



Fondo de Investigación y Desarrollo en Educación - FONIDE

Departamento de Estudios y Desarrollo.

División de Planificación y Presupuesto.

Ministerio de Educación.

Adherencia y dosis de los equipos de aula al programa Un Buen Comienzo y su relación con los resultados en los párvulos

Investigadora Principal: Susana Mendive C.

Investigadores Secundarios: Christina Weiland

Institución Adjudicataria: Pontificia Universidad Católica de Chile

Proyecto FONIDE N°: F611106_2011

Enero 2013

Información: Secretaría Técnica FONIDE. Departamento de Estudios y Desarrollo – DIPLAP. Alameda 1371, Piso 8, MINEDUC.
Fono: 406 6073. E-mail: fonide@mineduc.cl

Adherencia y dosis de los equipos de aula al programa Un Buen Comienzo y su relación con los resultados en los párvulos

INFORME COMPLETO

Antecedentes

Estudios diseñados para conocer el impacto de programas educativos y sociales, requieren el complemento de estudios de fidelidad de implementación (Fdl), que revelan el grado de correspondencia entre el programa propuesto y el efectivamente implementado (Century, Rudnick, & Freeman, 2010; Durlak, 2010). En el año 2007, la Asociación para la Educación de Niños Pequeños (NAEYC) y la Sociedad de Investigación en Desarrollo Infantil (SRCD) han declarado que “la investigación en implementación es una prioridad en educación para la primera infancia” (Durlak, 2010, p. 349). Diversas razones motivan una declaración como la citada. En primer lugar, el nivel de implementación informa la validez de los resultados de impacto de un programa. En ausencia de efectos, si no se conoce el grado de implementación de un programa, se corre el riesgo de cometer error de tipo III, o concluir que el programa no es efectivo, cuando en realidad el problema se debe a problemas de implementación del modelo esperado (Dusenbury, Brannigan, Falco, & Hansen, 2003). En cambio, si se cuenta con estudios de Fdl se puede establecer específicamente si la falta de efecto se debió a problemas de implementación o a problemas en la teoría de cambio (Eidelman, Weiland, & Yoshikawa, 2011).

En segundo lugar, conocer el nivel de implementación de un programa también entrega lecciones relevantes a considerar en el diseño de políticas públicas. Por un lado, informa acerca de su factibilidad. Un muy bajo nivel de implementación sugiere que es un programa no viable ni recomendable para expandirse a nivel de política pública (Dusenbury, et al., 2003). Por otro lado, los estudios de fidelidad de implementación permiten comprender qué mecanismos específicos explican los efectos alcanzados por un programa, y por lo tanto, informan en qué aspectos de este debe ponerse énfasis y recursos en un intento de expansión.

El potencial de los estudios de Fdl para conocer los mecanismos que explican los efectos de un programa, de especial relevancia para informar políticas públicas, solo puede satisfacerse si se utiliza una aproximación experimental y no solo descriptiva (Cordray & Pion, 2006; Cordray & Hulleman, 2009). La *aproximación experimental* en estudios de Fdl explora si existe diferencia en los mecanismos o componentes causales de un programa, por ejemplo, implementación de actividades de aprendizaje de calidad, entre el grupo que recibió la intervención y el grupo de comparación. La asignación al azar de las unidades de observación a los distintos grupos, y la comprobación que ambos grupos parten en similares condiciones en las variables de resultado antes de la intervención, permite concluir que las diferencias de resultados entre los grupos, se deben a la acción de los mecanismos causales del programa (Shadish, Cook, & Campbell, 2002). Sin embargo, es razonable suponer que en estudios experimentales llevados a cabo en “el terreno” (a diferencia de los que se llevan a cabo en contextos de laboratorio), los individuos del grupo de comparación también pueden recibir elementos del programa por otras vías (ej. educadoras del grupo de intervención cambiaron al de comparación, transferencia de conocimiento entre sujetos de ambos grupos, o educadoras del grupo de comparación han aprendido elementos del programa en otras instancias de desarrollo profesional). Por lo tanto, la comparación entre grupos de tratamiento se torna

relevante (Kerr, Kent, & Lam, 1985) y se ha propuesto como otra dimensión en los estudios de Fdl (Durlak, 2010).

En cambio, la *aproximación descriptiva* explora la relación entre dimensiones de implementación del programa y los resultados obtenidos solo en el grupo que recibió la intervención. Conocer que el nivel de implementación del programa predice resultados de los niños, si bien es relevante, no permite despejar la posible influencia de otros factores no controlados en un estudio descriptivo, y por lo tanto, no se puede concluir que los recursos invertidos en el programa fueron los que causaron los resultados obtenidos. Los estudios reportados a la fecha sobre Fdl y su relación con los resultados de programas en contextos educativos del nivel parvulario, se concentran en la aproximación descriptiva (e.g., Domitrovich, et al., 2010; Greenwood, Tapia, Abbott, & Walton, 2003; Hamre, et al., 2010; Odom, et al., 2010; Wasik, & Hindman, 2011) en vez de la experimental (cf. Darrow & Dickinson, 2011; Justice, Mashburn, Pence, & Wiggins, 2008).

En Chile, entre los años 2008 y 2011 se implementó el programa Un Buen Comienzo (UBC), un programa de desarrollo profesional docente (DPD) que busca mejorar el desarrollo del lenguaje y alfabetización inicial, desarrollo socioemocional, involucramiento familiar y coordinación del sistema de educación parvularia con los servicios de salud en Chile. UBC se dirigió a los niveles de transición 1 y 2 (pre kínder y kínder) de escuelas municipales que atienden a niños/as en condición de vulnerabilidad social en la Región Metropolitana. El programa se implementó en el contexto de un estudio experimental con el objetivo de conocer sus efectos en diversos aspectos del desarrollo infantil y en las prácticas docentes de educadoras y técnicos en párvulos (en adelante “equipos de aula”) (Yoshikawa, et al., 2012). Este estudio de gran magnitud y potencial para explicar efectos, marcará la historia del conocimiento acerca de efectividad de programas en el nivel de educación parvularia chilena, gracias a su diseño experimental, que es el tipo de estudios que suele estar ausente en la literatura de evaluación de efectividad de programas educativos chilenos (Bellei, Raczynski, & Osses, 2010). El presente estudio pretende informar políticas públicas en el nivel de educación parvularia, indagando si existen diferencias en los mecanismos causales centrales del programa UBC entre el grupo que recibió el programa y el grupo de comparación, y si estas diferencias tienen relación con el cambio en los resultados de los niños.

Para el caso de nuestro país, resulta particularmente importante conocer si un programa como UBC logra modificar las prácticas docentes e incrementar los niveles de desarrollo infantil debido a que a) comprensión oral, vocabulario, y escritura emergente, pilares de la intervención UBC en alfabetización inicial, son centrales para el desarrollo lector y académico posterior (e.g., Cooper, Roth, & Speece, 2002; Dickinson, 2011; National Early Literacy Panel, 2008), b) la instrucción explícita en vocabulario y estrategias de comprensión lectora no son parte de las prácticas docentes cotidianas en las aulas del nivel preescolar (Facultad de Educación, PUC, 2011; Strasser, Lissi & Silva, 2009) y c) precisamente vocabulario y comprensión oral son las habilidades más deficitarias en los niños de prekínder que viven en condición de vulnerabilidad socioeconómica en Chile (Yoshikawa et al., 2008).

Revisión completa de antecedentes

Fidelidad de implementación y dimensiones estudiadas. Fidelidad de implementación es un constructo multidimensional, que contempla entre sus principales dimensiones, adherencia, dosis, calidad, responsividad de los participantes y comparación entre grupos de intervención (Durlak, 2010). *Dosis*, o la cantidad implementada de un

programa, y *adherencia*, o el nivel de implementación de los elementos centrales sugeridos por los programas, son las dimensiones recurrentemente exploradas en estudios de FdI (Durlak, 2010), y en el caso de programas en educación parvularia han entregado importantes pistas para comprender efectos (Dickinson, 2011; Farran, et al., 2011; Hamre, et al., 2010).

Dosis de intervención: observación directa se relaciona con resultados de los niños. Dosis refiere a cuanto del programa propuesto es finalmente implementado (Durlak, 2010). Si un programa persigue cambios en los niños como variables de resultado, es de especial interés conocer a cuánto de los mecanismos causales del programa han sido expuestos los niños.

Respecto a los desafíos de medición de la variable dosis, existe poca investigación desarrollada, pero de la evidencia disponible, se puede concluir que es preferible recoger datos a partir de la observación directa de prácticas docentes y así construir una estimación de la dosis típica entregada del programa, en vez de recurrir al autorreporte de los docentes, el que estaría sesgado por deseabilidad social y no ha mostrado relación con resultados de programas (Hamre et al., 2010). Además, la observación directa de prácticas parece ser una muestra representativa del trabajo que cotidianamente se realiza en el aula, más aun cuando esta evaluación no está asociada a evaluación de desempeño con consecuencias laborales (Preiss, et al., 2011).

La limitada evidencia respecto a la relación entre dosis de exposición a experiencias de alfabetización inicial medida por observación directa, y los resultados de los niños, converge en una relación positiva. En el estudio citado de Hamre y colegas, del programa *My teacher partner-Language and Literacy Curriculum*, se operacionalizó la variable dosis a través del promedio de minutos observados en actividades que correspondían con el programa, solo en el grupo que recibió la intervención. De acuerdo a ello, los autores encontraron que a mayor extensión de tiempo de actividades correspondientes con el programa por parte de las educadoras, mayor fue el crecimiento de alfabetización inicial de los niños en un período de nueve meses. Similarmente Connor, Morrison y Slominski (2006), encontraron que a mayor tiempo en actividades focalizadas en el reconocimiento del código se asoció con el crecimiento en reconocimiento de letras y palabras de los niños, y mayor tiempo destinado a experiencias relacionadas con significados (por ejemplo lectura de cuentos), se asociaron con el crecimiento en vocabulario infantil durante prekinder. Por tratarse de evidencia chilena, de especial relevancia es el hallazgo de Strasser y Lissi (2009), quienes encontraron que aún una escasa exposición a experiencias de enseñanza explícitas ligadas al reconocimiento del código en kínder, mostró un efecto único en conocimiento del alfabeto y escritura emergente en el mismo año, así como en reconocimiento de letras y palabras medido con Woodcok Muñoz en primero básico, luego de controlar por nivel de desarrollo previo en las mismas dimensiones de los niños, nivel educacional de la madre, exposición a libros y enseñanza de letras en el hogar, en una muestra de 12 aulas de diversa de dependencia, nivel socioeconómico y calidad de los establecimientos.

En este estudio, se conceptualizará dosis como el tiempo total de tópicos abordados por el programa UBC al que los niños fueron expuestos por parte de sus educadoras.

Adherencia: observación directa se relaciona con resultados de los niños. Adherencia refiere a la implementación de los elementos centrales de un programa (Durlak, 2010). En el contexto de programas educacionales, los elementos centrales suelen ser prácticas pedagógicas consideradas de calidad que se transmiten a través de programas de desarrollo profesional docente, o de currículos específicos, siendo esta segunda, la modalidad con mayor evidencia disponible.

La aún limitada evidencia del rol de adherencia a prácticas pedagógicas de programas en resultados de desarrollo infantil, indica que sí existiría una relación positiva entre ambos. La evidencia más robusta por su diseño experimental, aunque en una muestra reducida, proviene de un estudio que comparó la adherencia a prácticas pedagógicas de estimulación de lenguaje prescrito por el *Language-Focused Curriculum*, observada en 14 salas de prekinder distribuidas aleatoriamente entre grupo de intervención y de comparación. Los investigadores encontraron que la adherencia promediada de tres momentos durante un año, aceleró el lenguaje expresivo pero solo de aquellos niños que asistieron regularmente a clases (Justice, et al., 2008). Un resultado similar de la influencia de mayor adherencia al currículo a favor del grupo de intervención sobre resultados en los niños, se observó en un estudio experimental de la implementación del currículo preescolar Building Blocks de matemáticas en 36 aulas distribuidas en grupo de intervención, y dos grupos de comparación (Clements & Sarama, 2008). Por otro lado, en un estudio longitudinal de cohortes secuenciales llevado a cabo en una escuela, se encontró que la adherencia de educadoras en kínder y primer año básico a prácticas de alfabetización inicial, propuestas en el contexto de reforma de escuelas básicas, se asoció con un crecimiento más acelerado en lectura silenciosa y fluidez lectora de los niños en segundo año básico, comparado con un grupo de salas en donde no se expuso a los niños a las mismas estrategias pedagógicas, también seguido por tres años. La adherencia en este caso consistió en el total de estrategias observadas de un número total enseñadas, a través de observaciones mensuales del primer año de intervención (Greenwood, et al., 2003).

Investigación reciente apoya la relevancia de medir la adherencia a elementos del programa tanto en el grupo que recibe la intervención como en el de comparación, pues no siempre se obtiene diferencias significativas en la adherencia de los elementos centrales de un programa (Darrow & Dickinson, 2011)

En el presente estudio, se conceptualizará adherencia, como el total de estrategias pedagógicas de UBC a la que los niños fueron expuestos por parte de sus educadoras.

Características de educadoras relacionadas con adherencia y dosis de implementación. Para el diseño de políticas públicas, es relevante conocer si existen distintos niveles de adherencia y dosis de implementación de algún programa según características profesionales de los participantes. Esto informa por ejemplo, si es conveniente expandir un programa en grupos particulares de educadoras que demuestran ser proclives a implementar programas determinados. Por otra parte, permite rediseñar programas en aquellos grupos que se muestran menos proclives a implementarlos con alta fidelidad. La investigación actual sobre efectividad de programas educacionales, también sugiere la necesidad de conocer en qué circunstancias y grupos particulares un programa logra efectos positivos (Justice, et al., 2008).

A la fecha existen pocos estudios que han reportado la relación entre características de desarrollo profesional y nivel de adherencia o dosis de implementación de programas educativos. Por ejemplo, en el caso del programa de desarrollo profesional REDI que contempla estrategias de lenguaje y apoyo emocional, se encontró que más años de educación formal de las educadoras predijo mayor implementación de las estrategias pedagógicas propuestas por el programa de acuerdo a la percepción de las facilitadoras (coaches), en cambio, los años de experiencia y horas de entrenamiento no mostraron relación alguna con el nivel de implementación de las estrategias (Domitrovich, et al., 2009). Por otra parte, se ha reportado que a menos años de experiencia, los profesores tienen mayor orientación a seguir métodos y por lo tanto adhieren en mayor medida estrategias propuestas por un programa (Rohrbach, Graham, & Hansen, 1993; en Stein et al., 2008).

Respecto al contexto chileno, no existen estudios que hayan explorado las variables de interés de esta investigación en el nivel de educación parvularia, pero sí se ha reportado que

profesores de menos años de edad, con formación profesional de mayor duración y experiencia de desarrollo profesional en servicio, tienen mejor nivel de desempeño profesional, de acuerdo a la evaluación nacional obligatoria (Docente Más) en el nivel de enseñanza básica (Valencia & Manzi, 2011).

Aún cuando no exista suficiente evidencia ni convergente, por las razones anteriormente mencionadas resulta relevante conocer si en la experiencia particular del programa UBC las características profesionales juegan un rol en el nivel de dosis y adherencia. Por este motivo, en este estudio también se explorará si características de las educadoras, tales como edad, años de experiencia docente, participación en dependencia particular (que podría considerarse una experiencia de desarrollo profesional de calidad); predicen adherencia y dosis al programa.

Cambio en adherencia y dosis: nueva vía para conocer la relación entre Fdl y los efectos de los programas. Tradicionalmente se ha explorado en forma concurrente la relación entre Fdl a programas educativos y los efectos obtenidos. Sin embargo, en el debate actual acerca del rol de Fdl y efectos de los programas se ha planteado la necesidad de realizar estudios que capturen el cambio entre fidelidad y resultados de los niños en al menos tres momentos. Estudios de esta naturaleza permitirían poner a prueba la hipótesis que la falta de relación encontrada entre fidelidad de implementación al programa y efectos en los niños, es que las observaciones no sean suficientemente frecuentes como para detectar covarianza entre las medidas de fidelidad y los efectos (Domitrovich, et al., 2010).

Desde el punto de vista de la política pública, explorar el cambio entre Fdl y resultados de programas también resulta de relevancia. Este tipo de diseños permite explorar la hipótesis que el nivel de implementación podría ser mayor en la medida que los equipos de sala llevan más tiempo de participación en el programa y se sienten más adaptados a él. Más específicamente, interesa conocer si los niveles de implementación varían en el tiempo dependiendo de características sociodemográficas de las educadoras. Una posibilidad es que educadoras con más años de servicio muestren mayor nivel de implementación antes que sus pares con menos años de servicio. Si así se demostrara con este estudio, ello podría sugerir que futuras iniciativas de desarrollo profesional docente serían más efectivas si ofrecieran estrategias específicas para ajustarse mejor a grupos de profesores que tienen menos años de experiencia en aula y que manifiestan mayores dificultades para adaptarse desde el comienzo.

Consecuentemente, el estudio abordó las preguntas de investigación (a) ¿Existe cambio en las variables de adherencia y dosis de UBC a la que han sido expuestos los niños en las áreas de intervención de alfabetización inicial, manejo de grupo y de salud? (b) ¿Las características de las educadoras (educación, edad, experiencia) predicen el crecimiento en adherencia y dosis de implementación de UBC?, y (c) ¿El cambio en las variables de adherencia y dosis de UBC explican el crecimiento de los niños en lenguaje, desarrollo socioemocional, y función ejecutiva, controlando por características de los niños y calidad de las oportunidades de aprendizaje ofrecidas por las educadoras? Para las tres preguntas, se explorará si existe un patrón diferente entre el grupo que recibió la intervención UBC y el de comparación.

Para responder estas preguntas, se observaron los videos de las salas en las que los niños de ambos grupos de intervención han participado en los tres momentos de evaluación, y se determinó tanto la adherencia como la dosis de los elementos centrales del programa para luego examinar su relación con los efectos en los párvulos, controlando por características de los equipos de aula y de los niños.

Método

Características del programa *UBC* y estudio de impacto

Características del Programa. Un Buen Comienzo es un programa de desarrollo profesional docente orientado a tres dominios de desarrollo infantil (lenguaje y alfabetización inicial, salud, y desarrollo socioemocional). Los principales mecanismos (cambios intermedios) a través de los cuales se esperó que la intervención ejerciera un efecto, fueron la calidad de la interacción educadora-niño y prácticas de salud en el aula y el hogar. El programa de desarrollo profesional UBC incluye en cada mes (con un total de 6 meses anuales durante dos años) una sesión de capacitación de aproximadamente cuatro horas, y dos visitas, de aproximadamente 90 minutos, de acompañamiento en aula a cargo de una facilitadora, basado en el modelo Content-Focused Coaching (Staub, West, & Bickel, 2003). Un equipo de implementación capacitado entregó el programa a los equipos de sala de seis comunas de la Región Metropolitana, agrupadas en tres cohortes entre los años 2008 y 2011.

El énfasis del programa fue promover el desarrollo de habilidades de alfabetización inicial y de lenguaje oral, capacitando a las educadoras en la utilización un enfoque integrado para desarrollar habilidades de comprensión oral, vocabulario y escritura emergente (Snow, Burns, and Griffin 1998). En cuanto al componente socioemocional, se entrenó a los equipos de sala para que incorporen estrategias de regulación de la participación, normas de convivencia, y formas de motivar el aprendizaje durante las experiencias de aprendizaje.

Diseño del estudio de impacto de UBC. Para evaluar el impacto del programa se diseñó un ensayo controlado aleatorio de conglomerados (cluster randomized controlled trial), longitudinal en tres momentos de medición. Vía sorteo público se asignó aleatoriamente a 64 escuelas de 6 comunas participantes en la Región Metropolitana a los grupos de intervención 1 (que recibió el programa UBC incluyendo un set de libros infantiles de calidad), e intervención 2 o grupo de comparación (que recibió un set de libros y un taller de autocuidado). La implementación y evaluación del programa se realizó en tres cohortes en períodos de dos años cada una (Figura 1).

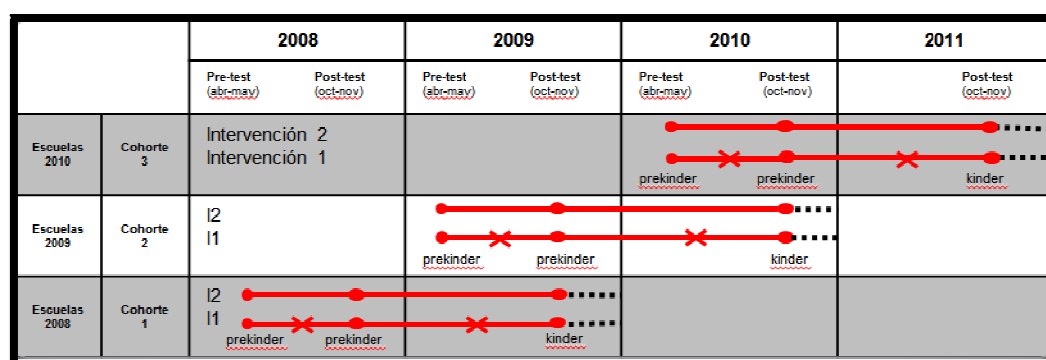


Figura 1. Relación entre Años y Cohortes en Estudio de Impacto UBC.

Participantes del estudio de impacto*¹. Un total de 64 escuelas, 140 educadoras, 110 técnicos y 1876 niños participaron en el estudio de UBC. Participación es definida como: ser reclutado, haber entregado consentimiento y ser aleatorizado. La tasa de reclutamiento fue de 85.9% y de respuesta 82.6% (la tasa de respuesta fue calculada como el número de evaluaciones completas respecto del número aproximado de familias elegibles basadas en la matrícula de la escuela). La tasa de reclutamiento de las familias fue el número de niños con consentimiento informado firmado respecto del número aproximado de familias elegibles basadas en la matrícula de la escuela. Para más información sobre la selección de las municipalidades, escuelas, educadoras, y niños, ver el Anexo 1.

Algunas educadoras se rehusaron a ser filmadas, lo que constituyó la fuente de información para evaluar tanto la calidad de las interacciones para el estudio de impacto, como las variables del presente estudio. En tiempo 1, 90 aulas fueron filmadas y observadas, 87 en tiempo 2 (de un total de 91 aulas para ambos tiempos), y 94 de 110 aulas en tiempo 3. El mayor número de aulas en tiempo 3 refleja el hecho que el estudio UBC siguió a los niños, no a las educadoras, entre los años de prekinder y kínder. En las escuelas suele haber mayor cantidad de salas en kínder que en prekinder. Por ejemplo, si los niños en el segundo año se distribuían en diversas salas de kínder, se realizaron filmaciones en todas las salas de kínder.

Procedimiento de estudio de impacto. La evaluación tanto de los niños, como de las prácticas docentes de sus equipos de aula, fueron obtenidas en tres momentos: pre-test (a inicios de NT1 o prekinder) y dos post-test (fines de prekinder y de NT2 o kínder) (Figura 1). Para evaluar la calidad de las prácticas docentes, se filmó una jornada normal en el aula (aproximadamente de 4 horas), de la cual siguiendo los criterios del instrumento CLASS (Pianta, La Paro, & Hamre, 2008), se seleccionaron cuatro segmentos de 20 minutos para evaluar la calidad de las prácticas de aula. El criterio para obtener los segmentos fue la existencia de un grupo de mínimo 5 niños y un adulto del equipo de aula a cargo, interactuando durante 20 minutos continuos.

En las escuelas asignadas al grupo de intervención UBC, participaron en el programa los equipos de aula de los niveles NT1 y NT2 en ambos años de la intervención. Para el estudio de impacto participaron solo los niños y educadoras que en el primer año estuvieron en NT1, y en el segundo año se siguió a los mismos niños en NT2 y se evaluó a sus educadoras (Tabla 1).

¹ Todas las secciones marcadas con el símbolo * fueron traducidas textual del reporte de estudio de impacto UBC (Yoshikawa et al., 2012)

Tabla 1.

Niveles participantes en capacitación y en estudio de impacto en escuelas asignadas al grupo de intervención UBC.

		Prekinder (NT1)	Kinder (NT2)
Año 1	Participa en Capacitación	Sí	Sí
	Participa en Estudio Impacto	Sí	No
Año 2	Participa en Capacitación	Sí	Sí
	Participa en Estudio Impacto	No	Sí

Diseño del presente estudio

Este estudio mantiene el diseño experimental y longitudinal del estudio de impacto UBC, así como la cantidad de escuelas, aulas, niños, y educadoras. Como principal fuente de información se utilizaron los videos de la evaluación CLASS. Asimismo, se consideraron los datos ya recogidos tanto de los párvulos como del cuestionario de desarrollo profesional docente.

Participantes. El estudio de impacto de UBC fue diseñado a fin de obtener un poder estadístico adecuado para detectar efectos pequeños de tratamiento. De hecho, a partir de resultados preliminares del estudio de UBC (Yoshikawa, et al., 2012), se ha visto el poder estadístico de este tamaño muestral para encontrar impactos. La muestra de este estudio de Fdl usa la muestra completa utilizada en el estudio de impacto, de modo que también es adecuada para detectar efectos pequeños de fidelidad en los resultados de los niños.

Según datos recolectados para tiempo 1 de observación, las educadoras tuvieron en promedio 44.5 años ($DS = 9.24$), un 58% ha enseñado por más de 15 años ($SD = 0.49$), un 18% ha enseñado menos de 5 años ($SD = 38\%$), un 59% ha tenido experiencia docente en un establecimiento privado ($SD = 0.49$), y un 45% ha tenido alguna formación de diplomado o postítulo ($SD = 0.49$). Como se dijo anteriormente, en estas variables los grupos de tratamiento fueron balanceados al inicio del estudio de impacto.

La continuidad de las educadoras entre los años 1 (medición en tiempo 1 y 2) y 2 (medición en tiempo 3) no fue un aspecto controlado por el presente estudio debido a que el estudio de impacto definió seguir a los niños y no a las educadoras. Para determinar si variables de las educadoras predijeron el crecimiento en adherencia y dosis, de la pregunta de investigación 2 del presente estudio, fue relevante conocer si la educadora se mantuvo en la misma sala entre tiempo 2 y tiempo 3.

De acuerdo al reporte de las codificadoras en la segunda fase de codificación de este estudio, 37 educadoras (39% del total de aulas en tiempo 3) se mantuvieron entre tiempo 2 y tiempo 3 (Tabla 2)². La información de 13 de las 37 aulas que se mantuvieron, se obtuvo al consultar registros de nombres de educadoras de base de datos de los estudios de impacto e implementación, pues problemas de audio o imagen no permitieron determinarlo a partir de la simple observación. En 15 casos no pudo determinarse si la educadora de tiempo 3 fue la

² Dada la alta complejidad de la evaluación de fase 2, se optó por solicitar solo una comparación de educadoras entre tiempos de observación. La comparación seleccionada se decidió a partir del conocimiento previo del proceso de implementación que el mayor cambio ocurriría entre tiempo 2 y tiempo 3, que entre tiempo 1 y tiempo 2.

misma que en tiempo 2 (ya sea porque fueron salas que no existieron en tiempo 2 ni tiempo 1, o porque salas de tiempo 2 no se mantuvieron en tiempo 3, o porque se trató de una nueva sala en tiempo 3).

Tabla 2.

Total de aulas en tiempo 3 que mantuvieron a la misma educadora de tiempo 2, según grupo de intervención.

	Intervención	Comparación	Total
No	22	20	42
Sí	23	14	37
No se puede determinar	6	9	15
Total			94

A partir del conocimiento de la experiencia de implementación del programa, existen diversas razones para explicar que la educadora del año 1 no se mantuviera en la misma sala en el año 2. Algunas escuelas mantuvieron a un grupo de educadoras en prekinder y a otro grupo en kínder, por lo que las educadoras no siguieron a los mismos niños entre los dos años. En otros casos se trató de rotación de educadoras por razones médicas, de contratación en otra escuela, o cambio a otros cargos dentro de la misma escuela. También ocurrió que algunas educadoras se mantuvieron entre ambos años en la misma escuela, pero en año 2 estuvieron en una sala diferente a la identificada por el estudio de impacto en el año 1 (por ejemplo, en año 1 estuvo en la denominada 010101 y en año 2 estuvo en la sala 010102). No existe información sistemática que permita conocer cuantas salas estuvieron en cada una de las razones mencionadas, e incluso es posible que existan otras razones no conocidas.

Selección de la muestra de videos. Se seleccionaron aleatoriamente tanto 2 segmentos de 20 minutos de los videos utilizados para la evaluación del CLASS, como 2 segmentos de 20 minutos continuos que no correspondieron a segmentos del CLASS entre todos los posibles segmentos de 20 minutos que restaban del tiempo filmado. Asimismo, si algún segmento seleccionado tuvo la mitad o más del tiempo con problemas de audio y video, o tiempo fuera de la sala, se seleccionaron al azar otros dos segmentos de 20 minutos de reserva, y se codificaron hasta completar el total mínimo de 80 minutos de datos válidos para disminuir el total de datos faltantes (o missings). La Tabla 3 muestra la distribución de segmentos codificados de las aulas de la muestra del estudio. En los análisis de la variable dosis, se controló por el número de minutos totales codificados en las aulas para remover el efecto de dichas diferencias en los resultados.

Tabla 3.

Número de aulas codificadas en cada tiempo de medición y número de aulas con 4, 5, o 6 segmentos codificados

Tiempo de medición	Aulas codificadas	Aulas con 4 segmentos codificados	Aulas con 5 segmentos codificados	Aulas con 6 segmentos codificados
1	90	43	20	27
2	87	41	19	27
3	94	36	25	33

Al utilizar en este estudio muestras de tiempo elegible y no elegible para el CLASS, y al seleccionar al azar ambos tipos de segmentos (CLASS y no CLASS) es posible generalizar la dosis y adherencia del trabajo pedagógico encontrado a la jornada completa filmada.

Procedimiento de codificación. Para garantizar la confiabilidad de los datos, las variables dosis y adherencia, se midieron en fases de codificación separadas y secuenciadas. Para hacer más eficiente el proceso de codificación, en la primera fase se midió dosis y en la segunda se midió adherencia. La revisión exhaustiva de los segmentos minuto a minuto, para obtener dosis en la primera fase, permitió identificar segmentos completos de tiempo ocurrido fuera de la sala que en la fase 2 de adherencia no fueron revisados nuevamente (pues en ellos sería imposible medirla).

Para el constructo dosis solo se consideraron los tópicos de lenguaje y alfabetización, que suelen tener mayor extensión de tiempo que manejo de grupo y aplicación de alcohol gel. Para el constructo adherencia se consideró estrategias de alfabetización, de manejo de grupo y de salud. Las estrategias de manejo de grupo de UBC fueron consideradas solo en el constructo adherencia, debido a que suelen ser de muy corta duración (tiempo inferior a un minuto) y por lo tanto tenía sentido medir solo si hubo presencia o ausencia de ellas. La adherencia de los equipos de aula al uso de alcohol gel (estrategia de salud del programa UBC) se incluyó en la fase de codificación 1 como una variable dicotómica (presencia o ausencia), pues era de fácil identificación y no ponía en riesgo la confiabilidad de la variable dosis de alfabetización. La Tabla 4 muestra los elementos clave del programa, constructos de Fdl asociados y en qué fase se codificaron.

Tabla 4.

Relación entre tópico del programa, constructo de Fdl evaluado y fase de codificación

Tópico del Programa	Constructo de Fdl	Fase de codificación
Alfabetización inicial	Dosis y	1
	Adherencia	2
Desarrollo socioemocional y Manejo de grupo	Adherencia	2
Salud – alcohol gel	Adherencia	1

En la fase de codificación de dosis se asignó al azar las aulas a las codificadoras. Respecto a la fase de adherencia, se organizó que la misma codificadora viera el video de tiempo 2 y 3 de la misma sala (la sala sí era asignada al azar), a fin de poder determinar si la educadora era la misma entre esos momentos (es decir si la educadora se mantuvo o no entre el primer año con prekinder y el segundo año con kínder).

Estrategia de codificación de dosis. Las categorías y códigos fueron adaptados de estudios previos conducidos en el nivel parvulario (Goodson, Layzer, Smith, & Rimdzius, 2004; Mascareño, 2012; Strasser, 2006) (detalles de códigos y categorías en Anexo 2).

En la lista de abajo, los códigos 1 a 9 refieren a tópicos Instruccionales de Lenguaje, los códigos 10-14 a Otros tópicos Instruccionales (no de lenguaje), y 15 a 16 refieren a momentos de rutina o No Instruccionales. Los códigos acompañados de la sigla **(UBC)** son los que fueron considerados tópicos de alfabetización de UBC.

Instruccionales de Lenguaje

1. Rutinas orales
2. Conversación acerca de un tema
3. Conocimiento de lo impreso **(UBC)**
4. Conciencia fonológica (aislada del proceso de escritura)
5. Vocabulario **(UBC)**
6. Escritura emergente **(UBC)**
7. Comprensión oral **(UBC)**
8. Adulto lee en voz alta **(UBC)**
9. Dibujar y pintar después de escuchar una lectura

Otros instruccionales

10. Matemática o ciencias
11. Arte o motricidad fina
12. Motricidad gruesa y grupos humanos
13. Juego libre
14. Instruccionales pero no focalizado en lenguaje y no se puede identificar una categoría específica.

No Instruccionales

15. Comida
16. Transición
17. Dato faltante porque los niños estuvieron fuera de la sala o hubo problemas de audio y video.

Utilizando el software de codificación de videos Videograph (versión 3.5.0.1)³ (Rimmele, 2004), se codificó por intervalos predeterminados de tiempo (Bakeman & Quera, 2011) de 1 minuto cada uno, unidad de tiempo también utilizada en un estudio en curso con una muestra de los videos de este estudio (Mascareño, Bosker, & Doolaard, 2010) y un estudio previo (Strasser & Lissi, 2009; Strasser, Lissi, & Silva, 2009), lo que favorecerá la comparación y discusión de resultados para el contexto chileno. Esto implica que el codificador observa un minuto, detiene la acción y determina la actividad predominante durante el intervalo de una lista de códigos mutuamente excluyentes pertenecientes a las tres categorías posibles (instruccionales de lenguaje, instruccional otros, no instruccional). No fue posible marcar un código instruccional y transición simultáneamente. En el mismo minuto los codificadores podían marcar un código de la categoría instruccional de lenguaje, instruccional otros, o comida (de la categoría no instruccional). Se repite el mismo procedimiento cada un minuto hasta completar los 20 minutos

³ Se produjo un manual de codificación para utilizar este programa, ver Anexo 3.

de cada segmento, y a lo largo de los 4 segmentos (ver Guía de variables y descripción de la Base de Datos producida en Anexo 4).

En las Tablas 5, y 6, se resume el número de minutos codificados y utilizados en el análisis por tiempo de medición.

Tabla 5.

Número total de minutos codificados en cada tiempo de medición

Tiempo de medición	Minutos codificados	Minutos codificados como PAV ^a	Minutos utilizados en el análisis
1	8,476	1,788	6,688
2	8,175	1,879	6,296
3	9,073	2,075	6,998
Total	25,724	5,742	19,982

^a PAV=niños fuera de la sala o con problemas de audio y video

Tabla 6.

Promedio y rango de minutos válidos analizados por tiempo de medición

Tiempo de medición	Total de aulas	Promedio minutos	DS	Rango
1	90	94.18	15.57	79-120
2	87	93.97	15.44	79-120
3	94	96.52	15.73	79-120

Estrategia de codificación de adherencia. En conjunto con actores clave del diseño del programa UBC, se generó una pauta de observación en la que se consensuaron tanto las categorías de observación como los indicadores (detalles de todos los indicadores y valores, organizados en las dimensiones, en Anexo 7). La pauta de observación contuvo más indicadores que los utilizados para responder las preguntas de investigación de este proyecto. Todos los indicadores utilizados tuvieron como alternativa de puntaje los valores 1 y 0, significando adherencia y no adherencia respectivamente. En la tabla 7 se presenta la lista de los indicadores utilizados para construir los puntajes de adherencia de las distintas dimensiones del programa, para responder las preguntas de investigación de este estudio.

Tabla 7.

Dimensiones e indicadores utilizados para construir la variable adherencia a UBC

Dimensión	Indicadores
Vocabulario (13 indicadores)	<p>La palabra enseñada proviene de un texto</p> <p>La palabra enseñada proviene de un texto y educadora lee en qué contexto se usó*</p> <p>La definición es accesible a la comprensión de los niños</p> <p>La definición es precisa</p> <p>La palabra enseñada es de nivel 3</p> <p>La palabra enseñada es de nivel 2⁴</p> <p>Se ubica la palabra en el muro de palabras (sin desafío)</p> <p>Se ubica la palabra en el muro de palabras con desafío*</p> <p>Los niños tienen oportunidad de verbalizar la palabra</p> <p>Utiliza una estrategia de interacción oral con la palabra (entre preguntas, razones, ejemplos, elecciones, motora)</p> <p>Utiliza más de una estrategia de interacción oral con la palabra*</p> <p>Se enseña el significado de la palabra aplicándolo a un contexto</p> <p>Se enseña el significado de la palabra aplicándolo a dos o más contextos*</p>
Exposición a textos (8 indicadores)	<p>Se menciona o lee el título del texto.</p> <p>Se menciona o lee el título, muestra en qué parte del texto se encuentra. O bien, invita a los niños a decir, leer, simular que leen o localicen el título*.</p> <p>Menciona o lee el autor.</p> <p>Menciona o lee el autor y explica qué significa. O BIEN, Menciona o lee el autor y muestra en qué parte del texto se encuentra*.</p> <p>Menciona o lee el autor, muestra en qué parte del texto se encuentra y explica qué significa**.</p> <p>Muestra imágenes sólo una vez.</p> <p>Muestra imágenes en más de una ocasión*.</p> <p>Exposición a literatura infantil (si el texto leído fue de una de las siguientes alternativas: texto literario, texto no literario).</p>
Comprensión oral (5 indicadores)	<p>Se utiliza una estrategia de comprensión oral (entre las opciones predecir, resumir, visualizar y hacer conexiones)</p> <p>Se utiliza la estrategia de comprensión oral más de una vez*</p> <p>Se utiliza más de una estrategia de comprensión oral (entre las opciones predecir, resumir, visualizar y hacer conexiones)</p> <p>Se anuncia la estrategia utilizada (entre las opciones predecir, resumir, visualizar y hacer conexiones)</p> <p>Se realizan preguntas de verificación de la comprensión del texto</p>
Escritura emergente (6 indicadores)	<p>El contenido de la escritura proviene de un texto o suceso ocurrido en el grupo</p> <p>El contenido de la escritura es explícitamente contextualizado por la educadora*</p> <p>Niño escoge la palabra a escribir entre opciones dadas</p> <p>Niño escoge el contenido de la escritura*</p> <p>Educadora apoya durante la escritura de los niños</p> <p>El apoyo durante la escritura desafía a los niños (incluye cualquiera de las siguientes: da el nombre de consonantes, da el sonido de letras, Indica al niño que busque la letra en algún recurso, alarga el sonido para que el niño lo segmente)*.</p>

⁴ En consulta con la experta que capacitó al equipo de implementación en las estrategias de vocabulario de UBC, se tomó la decisión de asignar un punto si la educadora selecciona una palabra de nivel 3 para enseñar a los niños, y 2 puntos para cuando seleccionan palabras de nivel 2. La racionalidad a la base fue que seleccionar palabras de nivel 3 es un intento de enseñar palabras que no son familiares para los niños, y por lo tanto preferible a palabras de uso frecuente o de nivel 1, pero no cumple cabalmente la propuesta de priorizar palabras de nivel 2 que fue enfatizado por el programa UBC.

Dimensión	Indicadores
Manejo de Grupo (17 indicadores)	El adulto realiza algún ejercicio físico breve en forma reactiva
	El adulto realiza algún ejercicio físico breve en forma proactiva*
	El adulto utiliza recurso para marcar inicio de actividad en forma reactiva
	El adulto utiliza recurso para marcar inicio de actividad en forma proactiva*
	El adulto utiliza recurso para regular al azar la participación en forma reactiva
	El adulto utiliza recurso para regular al azar la participación en forma proactiva*
	El adulto utiliza recurso para regular la conducta en forma reactiva
	El adulto utiliza recurso para regular la conducta en forma proactiva*
	El adulto asigna puesto o ubicación a los niños/as en forma reactiva
	El adulto asigna puesto o ubicación a los niños/as en forma proactiva*
	El adulto solicita que los niños levanten la mano para participar en forma reactiva
	El adulto solicita que los niños levanten la mano para participar en forma proactiva*
	El adulto asigna responsabilidades a un niño/a en forma reactiva
	El adulto asigna responsabilidades a un niño/a en forma proactiva*
	El adulto recuerda normas de convivencia en forma reactiva
	El adulto recuerda normas de convivencia en forma proactiva*
	El adulto otorga un espacio tranquilizador al niño/a disruptivo

Nota. Para dar más peso a indicadores que reflejaban mayor adherencia, si en el indicador marcado con * se obtuvo sí, obtuvo puntaje en el indicador y también obtuvo puntaje en el indicador inmediatamente anterior. Por ejemplo, si la educadora menciona o lee el autor y muestra en qué parte del texto se encuentra, obtuvo puntaje 1 en este indicador y también en el indicador 'Menciona o lee el autor', inmediatamente anterior. Además, si la educadora menciona o lee el autor, muestra en qué parte del texto se encuentra y explica qué significa, que tiene el símbolo **, obtuvo puntaje 1 en este indicador y a su vez 1 punto en los dos indicadores inmediatamente anteriores.

La Tabla 8 presenta el total y porcentaje del total de aulas que presentaron al menos un evento en los tópicos de alfabetización de UBC en el total de segmentos observados, respecto de los cuales se pudo codificar en qué medida lo observado en el evento adhería o no con la propuesta del programa UBC. Se aprecia alta variabilidad en el total de aulas que presentaron al menos un evento codificable. El tópico que en los tres momentos de medición presentó la menor cantidad de eventos fue comprensión oral con 18 a 30% del total de aulas de la muestra. Escritura emergente presentó alta variabilidad en la cantidad de aulas con eventos codificables entre los tiempos 1 y 3, a diferencia de exposición a textos cuyo rango varió solo entre 61 y 68% de aulas con al menos un evento codificable.

Debido a que el número real de aulas analizables disminuyó en forma significativa no fue posible crear un puntaje general de adherencia a alfabetización UBC (ya que se tendría un excesivo número de valor total de dato faltante), y no se tuvo suficiente poder estadístico para explorar si las características de las educadoras predijeron adherencia a estrategias de alfabetización UBC (parte de la pregunta de investigación 2), ni tampoco se pudo explorar la relación entre adherencia a estrategias de alfabetización UBC y resultados de los niños (parte de la pregunta de investigación 3.b).

Tabla 8.

Total de aulas y porcentaje del total con eventos codificables de adherencia en tópicos de alfabetización inicial, según tiempo de medición

Tiempo de medición	Aulas codificadas	Aulas con eventos de vocabulario	Aulas con eventos de exposición a textos	Aulas con eventos de comprensión oral	Aulas con eventos de escritura emergente
1	90	55 (61%)	55 (61%)	16 (18%)	49 (54%)
2	87	45 (52%)	54 (62%)	18 (21%)	67 (77%)
3	94	64 (68%)	63 (68%)	28 (30%)	76 (81%)

La codificación se llevó a cabo a través de una plataforma web que permitió ver los videos e ir registrando la evaluación en la misma pantalla; y la información de codificación se almacenó en una base de datos en forma automática. Se intentó igualar las condiciones de codificación y disminuir error de medición dotando de audífonos de buena calidad para las codificadoras.

Equipo de codificadoras y acuerdo entre observadores. La investigadora responsable y una ayudante estuvieron a cargo de la capacitación de las codificadoras. La ayudante es egresada de educación básica y educación de párvulos de la PUC. Además, previamente ella fue parte del equipo de codificadoras que en otro estudio observaron adherencia de facilitadoras y educadoras a UBC en las sesiones de acompañamiento (Mendive, Weiland, & Crosson, 2013). En la fase de codificación de dosis participaron 7 codificadoras (5 educadoras de párvulos en proceso de titulación, 1 educadora de párvulos con años de experiencia en sala, y una profesora básica). En la fase de codificación de adherencia se mantuvieron 4 de ellas y se agregaron 5 codificadoras más (2 educadoras de párvulos con experiencia en aula y 3 educadoras de párvulos en proceso de titulación). Para dosis y adherencia la capacitación constó de (a) Capacitación presencial (sesión de ocho horas); (b) Realización de dos prácticas de codificación con acompañamiento y reunión personalizada con la capacitadora; (c) Prueba de acuerdo (se realizaron dos pruebas con 4 videos de 4 segmentos cada uno hasta alcanzar 75% en todas las codificadoras y categorías). La ayudante y codificadoras se mantuvieron ciegas a los objetivos del estudio y a la pertenencia de las aulas a los grupos de intervención.

Medición de Acuerdo. Para la variable dosis, 10% de los videos seleccionados al azar, fueron también codificados por la ayudante (N=27). Para cada categoría se calculó un porcentaje de acuerdo para cada segmento de 20 minutos. Se obtuvo un porcentaje de acuerdo promedio de 89% para la categoría Instruccional Lenguaje (rango 83 a 97%), 86% en Instruccional Otros (rango 77 a 90%), 87% en No Instruccional (rango de 78 a 93%), 98% en problemas de audio y video (rango 96 a 100%).

Respecto a la variable adherencia, un 20% de los videos seleccionados al azar, fueron también codificados por la ayudante (N=55). En esta fase se midió por separado el acuerdo en la detección de eventos de las dimensiones observadas, y el acuerdo en los indicadores de las dimensiones, considerando solo los eventos que fueron detectados por codificadora y ayudante. Respecto a la detección de eventos, el porcentaje de acuerdo promedio fue de 67% en exposición a textos, 56% en vocabulario, 83% en comprensión oral, y 73% en escritura emergente (no se midió acuerdo en la detección de eventos en estrategias de manejo de grupo, pues se asumió que en todos los segmentos era posible evaluar esta dimensión). Todos los porcentajes de acuerdo obtenidos se consideran buenos, excepto el nivel medio alcanzado en vocabulario. La definición de eventos de vocabulario incluyó un amplio rango de conductas de la educadora que consideraron desde unas muy breves como mencionar un sinónimo de una palabra recientemente leída de un texto y retomar el relato de la historia, hasta destinar un tiempo más estructurado y prolongado de tiempo para explicar el significado de la palabra e involucrar a los niños en diversas formas de interactuar con el significado de la palabra nueva (ver página 4 de Anexo 7). El nivel medio de acuerdo obtenido en la detección de eventos de vocabulario, ocurrió principalmente en los eventos de corta duración y de carácter menos explícito, por lo tanto, se estima que de todos modos se debe haber alcanzado un alto nivel de acuerdo en los eventos de instrucción explícita y de mayor duración de vocabulario que son los propuestos por el programa UBC.

En cuanto al acuerdo en el nivel de los indicadores, se obtuvieron buenos porcentajes de acuerdo promedio, alcanzando 81% para exposición a textos (rango de 63 a 94%), 80% para vocabulario (rango de 73 a 97%), 95% para comprensión oral (93 a 100%), 85% para escritura (73 a 92%) y 87% para estrategias de manejo de grupo (77 a 91%). Una vez finalizado el proceso de codificación, se revisó el indicador “nivel de la palabra” (anexo 7), de la dimensión de vocabulario en el 100% de las veces que fue codificado, utilizando el registro de la palabra trabajada, más el registro de definición y de los contextos en los que la educadora aplicaba la palabra al explicarla a los niños. Esta revisión fue hecha en forma independiente tanto por la investigadora responsable de este estudio como por la persona experta que capacitó al equipo de implementación del programa en la estrategia de vocabulario, y luego se llegó a un consenso cuando existieron diferencias. Esta decisión responde a que este ítem se trató de un aspecto clave del programa UBC, pero requiere de un conocimiento avanzado.

En ambas fases, los resultados de la doble codificación se utilizaron para recalibrar a las codificadoras que tuvieran un porcentaje de acuerdo inferior al 70% a lo largo de las semanas de codificación. Como tomó más tiempo alcanzar acuerdo en la fase 2, se decidió tener un porcentaje mayor de doble codificación a fin de monitorear y retroalimentar los criterios de las codificadoras a lo largo del proceso.

Para ambas fases de codificación se cauteló que las comparaciones fueran sin previo aviso a las codificadoras y que todas tuvieran similar número de comparaciones a lo largo de las semanas de codificación.

Procedimiento de evaluación de los niños⁵. Las mediciones de inicio de año se llevaron a cabo entre Marzo y Mayo, y a fin de año se llevaron a cabo las evaluaciones entre Octubre y Diciembre. Para dar a los niños tiempo de adaptarse al contexto de aula, las mediciones comenzaron dos semanas después que comenzó el período escolar. Los datos de pretest se recolectaron antes de la aleatorización en las cohortes 2 y 3, y poco tiempo después de la aleatorización en cohorte 1 (8 de 64 escuelas estuvieron en cohorte 1). Las mediciones de los niños fueron conducidas en las escuelas durante una o dos sesiones de 30 a 50 minutos en un espacio seleccionado con ese fin fuera de la sala. Los evaluadores construyeron rapport con los niños y pasaron tiempo en el aula antes de evaluarlos individualmente.

Todos los evaluadores fueron egresados o titulados de psicología. Los evaluadores también condujeron el proceso de reclutamiento y consentimiento informado. Los evaluadores completaron entrenamiento intensivo, que incluyó información didáctica y administración de reglas para todas las mediciones, así como práctica administrando las evaluaciones (simulaciones con otros evaluadores y con niños de escuelas no participantes en la evaluación del estudio UBC). Para garantizar confiabilidad en la administración de las mediciones, un grupo de expertos dirigidos por miembros del equipo de Harvard, visitaron a los evaluadores durante el primer mes de recolección de cada tiempo de medición. Estos expertos entregaron retroalimentación continua a los evaluadores y sus supervisores.

Información de características sociodemográficas de educadoras, niños y familias. La información sociodemográfica de educadoras, niños, y familias fue recolectada solo una vez, generalmente al inicio del primer momento de medición. A las educadoras se les solicitó completar cuestionarios dentro de una semana. Los cuestionarios de las educadoras fueron recolectados en las escuelas. La edad de los niños y el género fueron obtenidos del reporte de los padres, también en cuestionarios.

⁵ Todas las secciones marcadas con el símbolo * fueron traducidas textual de Yoshikawa, et al., (2012), donde se reporta en detalle el estudio de impacto de UBC y las medidas de resultados de los niños.

Mediciones

Constructo dosis. Se definió *Dosis de alfabetización UBC* como el número total de minutos que la educadora destinó a tópicos específicos que eran foco del programa UBC (vocabulario, comprensión oral, escritura emergente, adulto lee en voz alta y conocimiento de lo impreso). La denominación dosis de UBC no pretende transmitir que estos tópicos son únicamente definidos ni promovidos por el programa UBC; más bien se denominó de esta manera para referirse en forma sintética el conjunto de tópicos que el programa promovió.

Constructo adherencia. Se definió adherencia como el promedio de los indicadores de cada tópico del programa: *adherencia a estrategias de manejo de grupo de UBC*, y para alfabetización, se obtuvo promedios separados para los tópicos vocabulario, exposición a textos, comprensión oral y escritura emergente.

VARIABLES DE RESULTADO DE LOS NIÑOS*

Lenguaje, alfabetización inicial, y matemáticas. En los tres tiempos de medición las habilidades de lenguaje y alfabetización inicial se midieron utilizando las siguientes subpruebas de la forma revisada en español de la Batería de Pruebas de Woodcock-Muñoz: vocabulario de imágenes, identificación de letras y palabras, dictado, comprensión de textos y problemas aplicados (o WMLS-R, Woodcock, Muñoz-Sandoval, Ruef, & Alvarado, 2005).

La subprueba de Vocabulario de Imágenes mide vocabulario receptivo en los primeros ítemes, requiriendo que el niño apunte las imágenes nombradas; y la principal proporción de ítemes mide vocabulario expresivo, en donde el niño debe nombrar la imagen mostrada. La subprueba de Identificación de Letras y Palabras es una medida de aprendizaje simbólico así como habilidades de decodificación lectora. Requiere que el niño conecte imágenes con palabras, nombre letras y lea en voz alta de una lista. La subprueba de Dictado mide habilidades de preescritura y escritura. En esta tarea, el niño escribe letras y palabras dictadas por el evaluador. La subprueba de Comprensión mide dicha capacidad para leer pasajes utilizando representaciones pictóricas para niños en un nivel preescolar.

Las habilidades matemáticas de los niños se midió utilizando la subprueba Problemas Aplicados de Woodcock-Muñoz Batería III Pruebas de Aprovechamiento, Forma en Español (Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005). La subprueba Problemas Aplicados es una medida de habilidades matemáticas emergentes, particularmente numeracidad y resolución de problemas simples. Las habilidades matemáticas no fueron parte del foco de intervención de UBC, de modo que esta subprueba se incluyó como una comparación de las habilidades objetivo.

Investigación previa ha demostrado altos niveles de confiabilidad interna, validez predictiva, y validez concurrente para estas cinco subpruebas (Alvarado, Ruef, & Schrank, 2005). Coeficientes de confiabilidad interna obtenidos se han ubicado en el rango de 0.76 a 0.97 dependiendo de la edad de los niños (Alvarado, Ruef, & Schrank, 2005). En todos los análisis, se utilizaron los puntajes brutos de cada subprueba.

Socio-Emocional. Comportamiento prosocial, externalizante, internalizante y control de impulsos – evaluaciones de padres y educadoras. Los padres evaluaron el Comportamiento Prosocial y Externalizante, mientras que las educadoras también evaluaron el Comportamiento Internalizante y Control de Impulsos. Ítemes de cada conducta fueron adaptados para Chile y fueron tomados de la Escala de Evaluación de Habilidades Sociales (Social Skills Rating Scale; SSRS; Gresham & Elliott, 1990) y del Instrumento de Desarrollo Temprano (Early Development Instrument EDI/ versión en español DIT; Janus & Offord, 2007).

Los padres evaluaron el comportamiento prosocial de los niños a través de los 6 siguientes ítems: “Se relaciona bien con los otros niños” “Tiende a dar, prestar, y compartir” (Gresham & Elliott, 1990); “Reconforta a un niño que está llorando o molesto,” “Se ofrece voluntariamente a limpiar un desorden que otro ha provocado”, “Juega y trabaja en forma cooperativa con los otros niños,” (Janus & Offord, 2007); y “Es colaborador con los otros” (Conduct Problems Prevention Research Group [CPPRG], 1990). Se condujo análisis factorial confirmatorio (AFC) separado por informante para el conjunto de ítems de Comportamiento Prosocial. Para cada informante, se encontró un constructo distinto. Tres ítems fueron eliminados debido a que las cargas factoriales fueron inferiores a 0.5. Para mantener consistencia a través de escalas e informantes, los mismos tres ítems fueron eliminados también de la escala reportada por la educadora. De acuerdo a los resultados del AFC, los 3 ítems que se mantuvieron de la escala de Comportamiento Prosocial utilizados en el análisis de este estudio incluyen: “Se relaciona bien con los otros niños”; “Juega y trabaja en forma cooperativa con los otros niños,” y “Es colaborador con los otros” Los alpha de Cronbach para la escala fueron de 0.65 y 0.64 en pretest y en el primer post test, respectivamente.

Los padres evaluaron las *conductas externalizantes* en 6 ítems: “Pelea con otros,” “Se siente irritable fácilmente,” “Tiene pataletas,” “Actúa impulsivamente.” (Gresham & Elliott, 1990); “Patea, muerde, y golpea a otros niños o adultos” (Janus & Offord, 2007); y “Quiebra cosas a propósito” (Werthamer-Larsson, Kellam, & Wheeler, 1991). Se condujo un AFC en la muestra completa de ítems de conducta externalizante. Se encontró apoyo a un constructo comprendido por los seis ítems. Los alpha de Cronbach para la escala fueron de 0.77 y 0.77 en pretest y posttest, respectivamente.

Todos los ítems de las escalas de comportamiento prosocial y externalizante fueron evaluados en una escala de 5 puntos, donde un alto puntaje correspondió a una conducta observada casi siempre, y un bajo puntaje a una conducta observada casi nunca. En análisis posteriores estas variables de resultado fueron promedios ponderados de las respuestas en los ítems de los distintos informantes identificados en el análisis factorial. Para obtener un puntaje en la variable resultado, cada niño requirió al menos dos tercios de los ítems posibles para no tener datos faltantes (missing), (excepto en Control de Impulsos, donde para cada niño se requirió al menos dos de los ítems posibles para que no hubiera dato faltante).

Control de impulsos y emoción positiva – medición del evaluador. Después de administrar la batería de medición de los niños, los evaluadores puntuaron 12 ítems reflejando la capacidad del niño para sostener la atención a las tareas, demostrar auto regulación, e involucrarse activamente para conseguir un objetivo, utilizando el Cuestionario de Orientación a la Tarea (Task Orientation Questionnaire o TOQ; adaptado de Smith-Donald, Raver, Hayes & Richardson, 2007). Cada ítem fue puntuado en una escala de 4 puntos, con claros descriptores conductuales para cada uno. Se condujo un AFC en el conjunto completo de ítems del cuestionario utilizando datos de los niños evaluados al inicio y fin de prekinder y se encontró tres constructos diferentes: emoción positiva, cambio en el objeto de atención, y control de impulsos.⁶ Los tres ítems para emoción positiva incluyeron si el niño “Se muestra alerta e interactivo; no es retraído,” “Demuestra placer al lograr hacer las tareas y al demostrar su habilidad para hacerlas activamente,” y “Se muestra seguro de sí mismo.” Los cuatro ítems para control de impulsos incluyeron “Puede esperar durante y entre tareas,” “Modula y regula

⁶ Los resultados difieren algo de los de Smith-Donald et al. (2007), quienes utilizaron análisis factorial exploratorio y encontraron apoyo para el factor de emoción positiva y un factor de cambio en foco de atención/ control de impulsos.

su excitación,” “Permanece en su silla durante las pruebas,” y “Es cuidadoso o destructivo con las materiales de las pruebas.” Los puntajes utilizados en el análisis fueron promedios ponderados de las respuestas de los niños a los ítems Emoción Positiva y Control de Impulsos. Para evitar puntaje faltante, cada niño tuvo que obtener al menos dos evaluaciones sin dato faltante para cada constructo. Los alpha de Cronbach para la escala de Emoción Positiva fueron de 0.68 y 0.70 para el pretest y postest, respectivamente; y los alpha de Cronbach para la escala de Control de Impulsos fueron de 0.83 y 0.82 para el pretest y postest, respectivamente.

Función Ejecutiva. La batería de pruebas incluyó medidas de dos de los componentes de función ejecutiva en el período de edad preescolar – cambio en el foco atencional y control inhibitorio. El primero es definido como la habilidad de dirigir la propia atención a un estímulo apropiado (Morrison, Ponitz, & McClelland, 2010). El segundo es definido como la habilidad de suprimir respuestas dominantes; los niños que exhiben control inhibitorio pueden detener respuestas y conductas automáticas y usar otras más apropiadas (Carlson & Moses, 2001; Diamond, Carlson, & Beck, 2005). Dentro de control inhibitorio, algunos investigadores distinguen entre dos sub factores: inhibición cognitiva e inhibición motora (McCabe, Rebello-Brito, Hernandez, & Brooks-Gunn, 2004; Nigg, 2000; Wiebe, Espy, Charak, 2008). La primera descansa en la habilidad de mantener pensamientos no deseados fuera de la mente para evitar actuar de acuerdo a la respuesta dominante, mientras que el segundo apunta a la capacidad inhibitoria del niño de controlar voluntariamente una respuesta primaria motora (Nigg, 2000). Tanto en el de impacto como en el presente estudio, se incluyeron medidas tanto de control inhibitorio cognitivo como motor.

Previo a los análisis de datos, se aplicaron dos reglas para remover la influencia de puntajes totales inválidos en las medidas directas de funciones ejecutivas. Los puntajes inválidos ocurrieron generalmente cuando un niño fallaba en los ensayos de práctica, cuando el evaluador no lo aplicó correctamente, o cuando el niño exhibió comportamientos que indicaban que la medición fue inválida. Primero, se estableció el puntaje total inválido como dato faltante, cuando hubo información que el puntaje total del niño se debió a alguna circunstancia de fatiga (ej. El niño no mostró lenguaje verbal, habló muy poco, lloró, rechazó al evaluador, se rehusó de participar durante la evaluación, estuvo tímido, enfermo, etc.). También se estableció como puntaje faltante cuando la medición fue interrumpida, o el evaluador cometió errores en la aplicación de la prueba. Segundo, se estableció puntaje total igual a cero, si el niño falló la práctica y no hubo información que indicara una situación de fatiga. También se asignó el puntaje total igual a cero para los niños cuyos evaluadores notaron que sus conductas violaban las reglas de la tarea consistentemente (ej. en la prueba golpe de lápiz, el niño no para de golpear el lápiz, imita o repite o lo que el evaluador hace, golpea continua o indiscriminadamente). Estos procedimientos fueron idénticos a aquellos utilizados en Bierman et al. (2008).

Cambio del foco de la atención. Se midió el cambio del foco de la atención a partir del Cuestionario de Orientación a la Tarea (Task Orientation Questionnaire o TOQ; adaptado de Leiter-R por Smith-Donald et al., 2007) respondido por los evaluadores. Los cuatro ítems de cambio en foco atencional incluyeron “Pone atención a las instrucciones y demostraciones”, “Es cuidadoso, interesado en dar las respuestas correctas”, “Sostiene la concentración – y se muestra dispuesto a hacer tareas repetitivas”, y “Coopera, cumple con los requisitos de los examinadores.” Alpha de Cronbach fue de 0.90 para las pruebas de pre y post test. Se utilizó un promedio ponderado de los cuatro ítems de cambio en el foco de atención, y se calcularon puntajes para cada niño que tuvo evaluaciones sin dato faltante en al menos dos ítems.

Control inhibitorio cognitivo. Se utilizaron dos medidas para evaluar el control inhibitorio cognitivo: la Clasificación de Tarjetas (Dimensional Change Card Sort o DCCS) y Golpe de Lápiz (Pencil Tapping o PT). En la prueba de Clasificación de Tarjetas (Zelazo, 2006; Frye,

Zelazo & Palfai, 1995), se les mostró a los niños tarjetas que variaron según las dimensiones de color y forma (ej. Rojo y azul; conejo y botes). Se les pidió a los niños ordenar las tarjetas de acuerdo a una dimensión (forma o color) y luego se les pidió ordenarlas de acuerdo a la otra dimensión. Para confirmar que los niños entendían las reglas, los evaluadores administraron ítems de práctica, seguido por 6 ensayos de la prueba. El puntaje final en DCCS representa el número de ensayos (de un total de 6) en que los niños tuvieron que inhibir el criterio previo y ordenar correctamente las tarjetas de acuerdo al nuevo criterio. El orden de las dimensiones evaluadas fue contrabalanceado (a la mitad de los niños se les pidió ordenar por forma primero y luego por color, mientras que a la otra mitad se le pidió ordenar por color y luego por forma).

En la prueba de Golpe de Lápiz (Diamond & Taylor, 1996), se solicitó al niño golpear dos veces si el evaluador golpeaba una vez, y una vez si el evaluador golpeaba dos veces. Los evaluadores primero administraron un conjunto de ensayos de práctica para asegurar que los niños comprendieron las reglas de la tarea. A los niños que pasaban la práctica, se les aplicaba los 16 ensayos totales. La tarea mide principalmente la habilidad de control inhibitorio cognitivo del niño y, en un grado menor, memoria operativa y actividad motora fina (Bierman et al, 2008). Los puntajes representaron el número correcto de ensayos de un total de 16.

En tiempo 1 no existieron diferencias significativas en variables de resultado de los niños ni de variables de su contexto relevantes, ni de las educadoras según condición de tratamiento, lo que permite concluir que ambos grupos resultaron balanceados luego de la asignación al azar a los grupos intervención y comparación (Yoshikawa, et al., 2012).

Covariables del niño (nivel individual). Se incluyeron covariables que controlaron por características de los niños. Ellas incluyeron edad del niño y género, utilizando datos obtenidos de los cuestionarios de los padres. La edad de los niños fue medida en meses desde el día de la evaluación en cada tiempo de medición. Respecto a género, en el tiempo de asignación aleatoria, se incluyó un indicador dicotómico si el niño era de género masculino.

Variables de las educadoras. Se examinaron características de las educadoras que son consideradas importantes en la literatura de interacciones de aula (Downer, Locasale-Crouch, Hamre, & Pianta, 2009), y también en el contexto sociodemográfico de Chile: edad de la educadora (en años), experiencia de enseñanza en establecimientos educacionales particulares pagados (predictor dicotómico), experiencia de enseñanza (en años), y educación de postítulo o diplomado (predictor dicotómico). Todas las variables de la educadora fueron obtenidas de los cuestionarios de desarrollo profesional docente. Si hubo más de una educadora, se promediaron las variables de las educadoras para crear un puntaje de nivel de aula. Respecto a calidad de las oportunidades de aprendizaje ofrecidas por las educadoras (considerada como variable de control de la pregunta de investigación 3), basado en las asociaciones encontradas con resultados de los niños en esta misma muestra al final del prekindergarten (Leyva, et al., 2013), se controló por calidad del apoyo instruccional medido a través del instrumento CLASS⁷.

Estrategia de análisis de datos

Se abordaron las preguntas de investigación utilizando dos estrategias de modelamiento: 1) modelos de regresión jerárquica lineal en tiempo 2 o 3, que controlan por una variable de resultado (ya sea dosis o una variable del niño) en tiempo 1; y 2) modelos de crecimiento longitudinal jerárquico. La estrategia (1) entrega resultados que facilita la interpretación respecto de patrones encontrados en cada tiempo de medición para la variable

⁷ Dado que del análisis factorial del CLASS se encontraron tres factores diferentes, no corresponde promediar los puntajes de todos los factores en un solo constructo para controlar por él.

de interés (o variable de resultado). La estrategia (2) permite examinar si hay cambio 'intra' aula o del niño en el tiempo, y si hay predictores de las diferencias en el cambio intra aula o intra del niño, en el tiempo a lo largo del período de tiempo completo del estudio (Singer & Willet, 2003).

Pregunta de investigación 1 - ¿Existe cambio en adherencia y dosis de implementación de UBC a la que han sido expuestos los niños en las áreas de intervención de alfabetización inicial, manejo de grupo y de salud?

Esta pregunta de investigación se subdivide en dos: (1.a) ¿Existe cambio en dosis de tópicos de alfabetización UBC a la que han sido expuestos los niños?; y (1.b) ¿Existe cambio en adherencia a UBC a la que han sido expuestos los niños en estrategias de alfabetización, manejo de grupo, y salud?

Se abordó la *pregunta de investigación 1.a* utilizando modelos separados en tiempo 2 y 3, y análisis de modelos de crecimiento longitudinal (conocido como Longitudinal Growth Model; LGM). Para los modelos de cada tiempo de medición, se especificó la siguiente ecuación,

$$Dosis_{ik} = \gamma_{00} + \gamma_{10}N_min_{ik} + \gamma_{11}TRATAMIENTO_k + \eta_k + \varepsilon_{ik} \quad (1)$$

Donde el subíndice i denota aulas individuales y k indica la escuela. Dosis será definida como la variable resultado o de interés, N_min es una variable de control que ajusta por el número total de minutos codificados en una sala, y TRATAMIENTO indica qué escuelas estuvieron en la condición de intervención versus comparación. η representa los interceptos aleatorios para las escuelas. ε es un término residual que representa la proporción del puntaje de la variable resultado del nivel aula, no predicha por el modelo hipotetizado.

Para los modelos de crecimiento longitudinal, se especificó la siguiente ecuación:

$$Dosis_{ijk} = \gamma_{00} + \gamma_{10}N_min_{ijk} + \gamma_{11}TRATAMIENTO_k + \gamma_{12}TIEMPO_{ijk} + (\eta_k + \gamma_{0ik} + \lambda_{1ik}TIEMPO_{ijk} + \varepsilon_{ijk}) \quad (2)$$

Donde todos los términos fueron definidos como en la ecuación (1). El subíndice j representa tiempo e identifica qué predictores variaron a lo largo de los momentos de medición. TIEMPO es un efecto fijo que representa el momento de medición (con valores 0, 1, y 2). γ es varianza no explicada de aulas en nivel de dosis inicial y λ es la varianza no explicada de la tasa de cambio.

Se abordó la pregunta de investigación 1.b utilizando modelos separados en tiempo 2 y 3. Para los modelos de cada tiempo de medición, se especificó la siguiente ecuación,

$$Adherencia_{ik} = \gamma_{00} + \gamma_{11}TRATAMIENTO_k + \eta_k + \varepsilon_{ik} \quad (3)$$

Donde el subíndice i denota aulas individuales y k indica la escuela. Adherencia a cada uno de los tópicos del programa (vocabulario, exposición a textos, comprensión oral, escritura emergente, y manejo de grupo) será definida como la variable resultado o de interés, y TRATAMIENTO indica qué escuelas estuvieron en la condición de intervención versus comparación. η representa los interceptos aleatorios para las escuelas. ε es un término residual que representa la proporción del puntaje de la variable resultado del nivel aula, no predicha por el modelo hipotetizado.

Para adherencia a estrategias de manejo de grupo, se especificó la ecuación (3) utilizando modelos jerárquicos lineales (conocido como Hierarchical Linear Model; HLM). Para adherencia a cada uno de los tópicos de alfabetización UBC, se usó análisis de regresión Tobit, en el que se modeló simultáneamente la diferencia, según grupo de tratamiento, en la probabilidad que se implemente una experiencia de aprendizaje del tópico, y en la probabilidad que esta sea implementada de la forma propuesta por el programa. Esta técnica de modelamiento fue necesaria para separar *dosís* (¿Ocurrió la implementación de un tópico?) y *adherencia* (si se implementó el tópico, ¿Fue de acuerdo a lo propuesto por UBC?) (McDonald & Moffitt, 1980). Debido a que un bajo número de aulas implementaron escritura, vocabulario, exposición a textos y comprensión oral (ver Tabla 8), no fue posible crear un puntaje general de adherencia a alfabetización UBC, ni tampoco se pudo explorar la relación entre adherencia a estrategias de alfabetización UBC y resultados de los niños (pregunta de investigación 3.b) debido a la significativa reducción de la muestra analizable, y la consecuente pérdida de poder estadístico necesario para conducir dichos análisis.

Para los modelos de crecimiento longitudinal de adherencia a estrategias de manejo de grupo UBC, se especificó la siguiente ecuación:

$$Adherencia_{ijk} = \gamma_{00} + \gamma_{11}TRATAMIENTO_k + \gamma_{12}TIEMPO_{ijk} + (\eta_k + \gamma_{0ik} + \lambda_{1ik}TIEMPO_{ijk} + \varepsilon_{ijk}) \quad (4)$$

Donde todos los términos fueron definidos como en la ecuación (2). El subíndice j representa tiempo e identifica que predictores variaron a lo largo de los momentos de medición. Y TIEMPO es un efecto fijo que representa el momento de medición (con valores 0, 1, y 2). γ es varianza no explicada de aulas en nivel de adherencia inicial y λ es la varianza no explicada de la tasa de cambio.

Pregunta de investigación 2 - ¿Las características de las educadoras (educación, edad, experiencia) predicen el crecimiento en adherencia y dosis de implementación de UBC?

Para abordar esta pregunta se agrega el siguiente término en la ecuación (1): $\gamma_{13}X_{ik}$, donde X es un vector de características de las educadoras y i denota aulas. Específicamente, se examinó si la educadora tiene estudios de postítulo o diplomado, su edad en años, y años de experiencia (en las categorías menos de 5 años, 5-14 años y más de 14 años).

Pregunta de investigación 3 - ¿El cambio en adherencia y dosis de implementación de UBC explican el crecimiento de los niños en lenguaje, desarrollo socioemocional, y función ejecutiva, controlando por características de los niños y de calidad general de las educadoras?

Esta pregunta de investigación se subdivide en dos (3.a) ¿El cambio en dosis de alfabetización de UBC explica el crecimiento de los niños en lenguaje y matemáticas?; y dada la pérdida de poder estadístico, la pregunta (3.b) que pudo ser respondida en este informe es ¿El cambio en adherencia a estrategias UBC de manejo de grupo, explican el crecimiento de los niños en el desarrollo socioemocional?

Para responder la pregunta de investigación 3.a, se especificó la siguiente ecuación en cada tiempo de medición:

$$Z_{lik} = \gamma_{00} + \gamma_{10}N_{-min}_{ik} + \gamma_{13}X_{ik} + \gamma_{14}Nino_{lik} + \gamma_{15}Dosis_k + \eta_k + \varepsilon_{lik} \quad (3)$$

Donde los términos fueron definidos como en las ecuaciones anteriores, con la adición de la variable de resultado Z de los niños (nivel individual).

Para los modelos de crecimiento longitudinal, la ecuación para la pregunta de investigación (3.a) fue:

$$Z_{ijk} = \gamma_{00} + \gamma_{10}N_{-min_{ikj}} + \gamma_{13}X_{ik} + \gamma_{14}Nino_{lik} + \gamma_{15}Dosis_{ikj} + (\eta_k + \gamma_{0lik} + \lambda_{1lik}TIEMPO_{lijk} + \varepsilon_{lijk}) \quad (4),$$

Donde los términos fueron definidos en las ecuaciones previas, pero γ es la varianza no explicada en el momento inicial para cada variable resultado de los niños, y λ es la varianza no explicada en la tasa de cambio de las variables de resultado de los niños.

Para responder la pregunta de investigación 3.b, respecto a adherencia a manejo de grupo, se especificó la siguiente ecuación en cada tiempo de medición:

$$Z_{lik} = \gamma_{00} + \gamma_{13}X_{ik} + \gamma_{14}Nino_{lik} + \gamma_{15}Adherencia_{ik} + \eta_k + \varepsilon_{lik} \quad (5)$$

Donde los términos fueron definidos como en las ecuaciones anteriores, con la adición de la variable de resultado Z de los niños (nivel individual).

Para los modelos de crecimiento longitudinal, se especificó la siguiente ecuación:

$$Z_{ijk} = \gamma_{00} + \gamma_{13}X_{ik} + \gamma_{14}Nino_{lik} + \gamma_{15}Adherencia_{ikj} + (\eta_k + \gamma_{0lik} + \lambda_{1lik}TIEMPO_{lijk} + \varepsilon_{lijk}) \quad (6)$$

Donde todos los términos fueron definidos como en la ecuación (4). El subíndice j representa tiempo e identifica que predictores variaron a lo largo de los momentos de medición. Y TIEMPO es un efecto fijo que representa el momento de medición (con valores 0, 1, y 2). γ es varianza no explicada de aulas en nivel de dosis inicial y λ es la varianza no explicada de la tasa de cambio.

Cabe aclarar que para la pregunta de investigación 3 no se presentan diferencias en los resultados de los niños según pertenencia al grupo de comparación o tratamiento (preguntas 3.a y 3.b), debido a que no tenemos permiso del equipo del programa. Estos resultados aun no son públicos. Una vez públicos se podrá incluir la variable de grupo de tratamiento en los análisis de resultados de los niños.

Resultados

Pregunta de investigación 1.a - ¿Existe cambio en dosis de tópicos de alfabetización UBC a la que han sido expuestos los niños?

Resultados de cada tiempo de medición. La tabla 9 presenta los estadísticos descriptivos para el total de tiempo destinado en cada variable, ponderando por número de minutos codificados en cada sala. Como se aprecia en la tabla, las aulas destinan poco tiempo en los tópicos de alfabetización UBC en tiempo 1 – 3.81 minutos en salas del grupo de comparación y 4.24 minutos en salas del grupo de intervención. En tiempo 2, las salas del grupo de comparación destinaron 7.63 minutos en tópicos de alfabetización UBC y aulas del grupo de tratamiento, 12.25 minutos. En tiempo 3, las salas del grupo de comparación destinaron 9.03 minutos en tópicos de alfabetización UBC, y las salas del grupo de intervención, 12.63 minutos. La Figura 2 entrega una referencia visual para estos hallazgos descriptivos.

Tabla 9.

Minutos promedio destinados a dosis de alfabetización UBC

Tiempo de medición	Tratamiento				Control			
	Promedio	DS	Min	Max	Promedio	DS	Min	Max
1	4.24	7.99	0	39	3.81	6.16	0	23
2	12.25	10.82	0	49	7.63	7.66	0	26
3	12.63	11.23	0	52	9.03	9.22	0	41

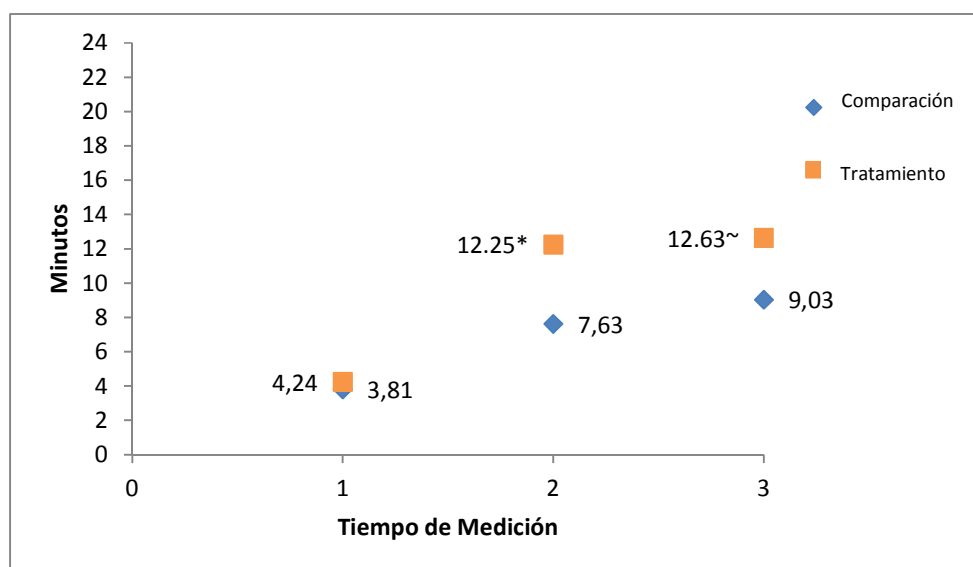


Figura 2. Promedio de Minutos de Dosis de Alfabetización UBC al Inicio de Prekinder, Fin de Prekinder, y Fin de Kinder.

~ $p < 0.10$. * $p < 0.05$.

En la Tabla 10, se presentan los resultados para los modelos de regresión de cada tiempo de medición (ecuación 1). La diferencia en dosis de alfabetización UBC por condición de tratamiento no fue estadísticamente significativa en Tiempo 1 ($p=0.78$), controlando por número de minutos codificados y utilizando efectos aleatorios por escuelas. En Tiempo 2, las aulas del grupo de intervención destinaron 5.22 más minutos en tópicos de alfabetización UBC que las aulas del grupo de comparación y la diferencia fue estadísticamente significativa ($p < 0.05$). En Tiempo 3, las aulas de intervención destinaron 3.83 más minutos en tópicos de alfabetización UBC que las aulas de comparación, y la diferencia fue estadísticamente significativa en un nivel de tendencia ($p=0.07$).

Debido a que el promedio de minutos codificados fue de 94 de un total de 218 minutos filmados (43%), en Tiempo 2, 5.2 minutos representa en una jornada diaria alrededor de 12.1 minutos más de Dosis de tópicos de alfabetización UBC en aulas del grupo de tratamiento en relación con el grupo de comparación. En Tiempo 3, se codificó 97 minutos en promedio de un total de 227 (42%), por lo tanto, el número comparable fue de 8.96 minutos más de dosis de

alfabetización de UBC en las salas del grupo de tratamiento en relación a las del grupo de comparación.

Tabla 10.

Análisis de regresión, diferencias en dosis de alfabetización UBC entre grupos de tratamiento y comparación, según tiempos de medición

	Tiempo 1 Dosis de alfabetización UBC	Tiempo 2 Dosis de alfabetización UBC	Tiempo 3 Dosis de alfabetización UBC
TRATAMIENTO	0.432 (1.556)	5.217* (2.176)	3.834~ (2.136)
N minutos codificados	0.143 (0.492)	-0.050 (0.067)	0.071 (0.068)
Intercepto	2.502 (4.729)	11.886~ (6.416)	2.240 (6.607)
Observaciones	90	87	94
CIC	0.000	0.115	0.000

Nota. Errores estándar en paréntesis. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ~ $p < 0.1$. CIC = Correlación Intra Clase.

Resultados de modelos de crecimiento longitudinal. También se utilizaron modelos de crecimiento longitudinal jerárquicos para examinar si hubo crecimiento en el tiempo de la dosis de tópicos de alfabetización UBC, y si la trayectoria de crecimiento de dosis UBC difirió por condición de tratamiento. Debido a que UBC siguió a los niños, y no a las mismas educadoras entre prekinder y kínder, es importante considerar que estos resultados deben ser interpretados como cambio en el tiempo de exposición a los niños a UBC.

Debido a que las escuelas fueron asignadas aleatoriamente a las condiciones de tratamiento después de la medición de Tiempo 1, y a que los modelos de corte transversal mostraron que no hubo diferencias significativas en dosis de alfabetización UBC en Tiempo 1 según condición de tratamiento, en Tiempo 1 se fijó el mismo intercepto para las salas de ambos grupos. Como se muestra en la Tabla 11, se encontró que hubo un crecimiento estadísticamente significativo en la dosis de exposición a los niños a alfabetización UBC (Modelo 1, $Tiempo=3.588$; $p < 0.001$). Además, hubo una pendiente más pronunciada en la trayectoria de crecimiento de dosis de alfabetización UBC a la que fueron expuestos los niños de las salas del grupo de intervención vs. de las salas del grupo de comparación (Modelo 2, $TRATAMIENTO * Tiempo=2.62$, $p < .01$). En términos prácticos estos resultados indican que la asignación al grupo de tratamiento incrementó la tasa de cambio de la dosis de alfabetización UBC desde el inicio de prekinder hasta el fin de kínder (ver Figura 3).

Tabla 11.

Modelos de crecimiento longitudinal de dosis de alfabetización UBC

	Modelo 1: Solo Tiempo	Modelo 2: Modelo Final
Tiempo	3.588*** (0.703)	2.133* (0.874)
N_minutos		0.014 (0.036)

TRATAMIENTO*Tiempo		2.616**
		(0.988)
Intercepto	4.916***	3.577
	(0.813)	(3.425)
Nivel 2, Varianza Escuela	0.000***	0.000***
	(0.00)	(0.00)
Nivel 2, Varianza en Tasa de cambio	10.371***	8.559***
	(4.289)	(3.749)
Nivel 2, Varianza en status inicial	0.000***	0.000***
	(0.00)	(0.00)
Nivel 1, Varianza Residual de Aulas	70.894***	70.945***
	(10.755)	(7.973)
N	271	271

Nota. Los modelos fueron calculados controlando por número de minutos codificados y utilizando efectos aleatorios de escuelas. Errores estándar en paréntesis. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ~ $p < 0.1$.

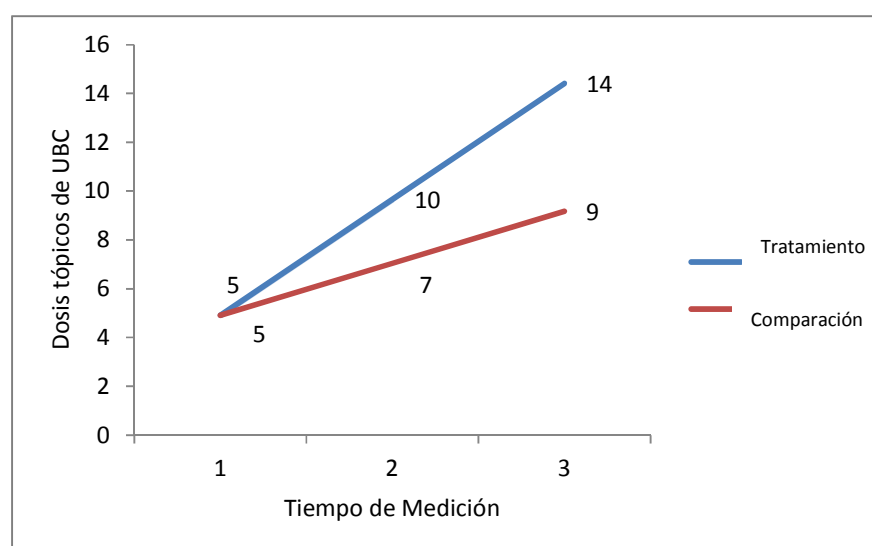


Figura 3. Crecimiento en dosis de alfabetización UBC en grupos de aulas de Tratamiento vs. Comparación desde el inicio de Prekinder al Fin de Kinder (controlando por número de minutos codificados)

Pregunta de investigación 1.b - ¿Existe cambio en adherencia a UBC a la que han sido expuestos los niños en estrategias de alfabetización, manejo de grupo, y salud?

Resultados de cada tiempo de medición. La Tabla 12 presenta los estadísticos descriptivos para el promedio de adherencia en cada tópico del programa UBC. Como se aprecia en la tabla, la adherencia en tiempo 1 fue más similar entre los grupos de tratamiento y comparación que en tiempos 2 y 3. Las aulas del grupo de tratamiento mostraron mayor adherencia a UBC en tiempos 2 y 3, comparado con tiempo 1, tanto en los promedios

obtenidos como en la cantidad de eventos de cada tópico de alfabetización inicial. Analizando solo los tópicos de alfabetización inicial, en tiempo 2 y 3 se aprecia que las educadoras adhirieron más a las estrategias de comprensión oral y escritura emergente; y adhirieron menos a las estrategias de exposición a textos.

Tabla 12.

Adherencia promedio por tópico específico del programa UBC y tiempo de medición

	Tratamiento					Comparación				
	N	Promedio	DS	Min	Max	N	Promedio	DS	Min	Max
<i>Tiempo 1</i>										
Vocabulario	31	0.23	0.13	0.00	0.46	24	0.23	0.12	0.00	0.54
Exposición a textos	30	0.16	0.17	0.00	0.50	22	0.09	0.20	0.00	0.63
Comprensión oral	11	0.35	0.31	0.00	1.00	5	0.24	0.09	0.2	0.40
Escritura emergente	30	0.22	0.17	0.00	0.6	17	0.21	0.12	0.00	0.42
Manejo de grupo	54	0.11	0.06	0.01	0.26	36	0.09	0.06	0.00	0.28
Uso alcohol gel	54	0.00	0.00	0.00	0.00	36	0.08	0.50	0.00	3
<i>Tiempo 2</i>										
Vocabulario	31	0.33	0.23	0.00	0.85	14	0.23	0.11	0.00	0.36
Exposición a textos	36	0.25	0.25	0.00	0.69	17	0.04	0.09	0.00	0.25
Comprensión oral	15	0.46	0.27	0.20	1.00	3	0.20	0.00	0.2	0.2
Escritura emergente	43	0.35	0.16	0.00	0.67	23	0.23	0.17	0.00	0.5
Manejo de grupo	51	0.13	0.07	0.01	0.28	36	0.09	0.07	0.00	0.26
Uso alcohol gel	51	0.33	0.97	0.00	6	36	0.00	0.00	0.00	0.00
<i>Tiempo 3</i>										
Vocabulario	33	0.29	0.17	0.00	0.77	31	0.22	0.11	0.00	0.46
Exposición a textos	33	0.16	0.21	0.00	0.63	28	0.11	0.17	0.00	0.63
Comprensión oral	15	0.41	0.29	0.00	0.80	13	0.28	0.13	0.20	0.60
Escritura emergente	43	0.34	0.24	0.00	0.83	33	0.27	0.18	0.00	0.67
Manejo de grupo	51	0.12	0.07	0.00	0.32	43	0.11	0.06	0.00	0.26
Uso alcohol gel	51	0.59	1.33	0.00	5	43	0.02	0.15	0.00	1

En la Tabla 13 se presentan los resultados de regresión utilizando análisis Tobit para los cuatro tópicos de alfabetización inicial, diferenciando en cada columna los resultados para cada tiempo de medición. Cuando en el 10% o más de las aulas la experiencia de aprendizaje fue cortada (porque el segmento llegó al minuto 20 y la actividad no pudo terminar de verse, o bien porque en el primer minuto del segmento la actividad ya había comenzado), se controló o removió el posible efecto que esta variable pudiera tener en la estimación de los parámetros. Como se aprecia en la significancia estadística del parámetro 'tratamiento' de cada tópico, se aprecia que solo en tiempo 2 las aulas que recibieron el programa, mostraron mayor

probabilidad de implementar experiencias de aprendizaje de los tópicos de alfabetización de UBC, y cuando las implementaban, también mostraron mayor probabilidad de hacerlo en forma consistente con la propuesta del programa (en el caso de exposición a textos la significación estadística fue en un nivel de tendencia $p < 0.10$).

De acuerdo a los resultados de los efectos marginales del tratamiento, las aulas que recibieron UBC mostraron una probabilidad de implementar eventos de vocabulario en 20 puntos porcentuales más que las aulas del grupo de comparación (dosis). Asimismo, las aulas que recibieron UBC tuvieron 0.20 puntos más en nuestra medida de adherencia a estrategias de vocabulario UBC. Esto es traducible a un tamaño efecto de 2,09. Para el caso de exposición a textos las aulas que recibieron UBC mostraron una probabilidad de implementar estos eventos de 18 puntos porcentuales más que las aulas del grupo de comparación (dosis); y obtuvieron 0.22 puntos más en la medida de adherencia a estrategias de exposición a textos de UBC, lo que equivale a un tamaño efecto de 2,44; ambos índices fueron significativos en un nivel de tendencia ($p < 0.10$). En el caso de comprensión oral, las aulas que recibieron el programa tuvieron una probabilidad de implementar estos eventos en 26 puntos porcentuales más que las aulas del grupo de comparación; y obtuvieron 0,39 puntos porcentuales más en la medida de adherencia a estrategias de comprensión oral UBC; este valor no puede traducirse a tamaño efecto debido a que la desviación estándar en tiempo 2 del grupo de comparación en comprensión oral fue igual a 0. Finalmente, se aprecia que el coeficiente de dosis de escritura emergente no fue significativo pero sí lo fue el de adherencia. Esto quiere decir que la asignación al grupo de intervención UBC no ejerció un efecto en la probabilidad que las aulas implementen experiencias de aprendizaje de escritura, pero cuando lo hacían, las aulas del grupo que recibió el programa tendieron a implementarlas del modo como UBC lo proponía, obteniendo 0,27 puntos más en la medida de adherencia que las aulas del grupo de comparación.

Desde un punto de vista práctico la magnitud de la diferencia entre ambos grupos para todos los tópicos de alfabetización es 'grande' de acuerdo a Cohen (1988) (ver Figura 4).

Tabla 13.

Estimación de parámetros y errores estándar de los análisis de regresión Tobit para dimensiones de alfabetización UBC, por tiempo de medición

	Tiempo 1	Tiempo 2	Tiempo 3
<u>Vocabulario</u>			
Tratamiento	-0.19 [0.19]	0.563* [0.25]	-0.086 [0.17]
<i>Efectos marginales del tratamiento</i>			
Probabilidad de aparición de algún evento de vocabulario (dosis)	-0.07 [0.07]	0.19* [0.09]	-0.03 [0.06]
Diferencia de grupos en adherencia, si algún evento de vocabulario ocurrió	-0.09 [0.10]	0.23* [0.10]	-0.05 [0.10]
Observaciones	90	87	94
<u>Exposición a Textos</u>			
Tratamiento	-0.064 [0.204]	0.436+ [0.222]	-0.084 [0.181]
Actividad cortada		0.84*** [0.19]	0.79*** [0.13]
<i>Efectos marginales del tratamiento</i>			
Probabilidad de aparición de algún evento de exposición a textos (dosis)	-0.03 [0.07]	0.18+ [0.09]	-0.03 [0.71]
Diferencia de grupos en adherencia, si algún evento de exposición a textos ocurrió	-0.05 [0.10]	0.22+ [0.12]	-0.05 [0.10]
Observaciones	90	87	94
<u>Comprensión oral</u>			
Tratamiento	0.522 [0.582]	1.781** [0.635]	-0.234 [0.345]
Actividad cortada			2.390** [0.322]
<i>Efectos marginales del tratamiento</i>			
Probabilidad de aparición de algún evento de comprensión oral (dosis)	0.07 [0.07]	0.26** [0.09]	-0.06 [0.09]
Diferencia de grupos en adherencia, si algún evento de comprensión oral ocurrió	0.11 [0.12]	0.39** [0.13]	-0.06 [0.09]
Observaciones	90	87	94
<u>Escritura emergente</u>			
Tratamiento	0.20 [0.24]	0.39* [0.16]	0.03 [0.14]
Actividad cortada	1.38*** [0.26]	0.72*** [0.21]	0.82*** [0.18]
<i>Efectos marginales del tratamiento</i>			
Probabilidad de aparición de algún evento de escritura (dosis)	0.08 [0.09]	0.10 [0.06]	0.01 [0.02]
Diferencia de grupos en adherencia, si algún evento de escritura ocurrió	0.08 [0.1]	0.27** [0.11]	0.03 [0.12]
Observaciones	90	87	94

Nota. Errores estándar en paréntesis, y son robustos a la agrupación de las aulas en las escuelas. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

En la Tabla 15, se presentan los resultados para los modelos de regresión de cada tiempo de medición (ecuación 3) respecto a adherencia a estrategias de manejo de grupo UBC. La diferencia en adherencia a estrategias de manejo de grupo de UBC por condición de tratamiento no fue estadísticamente significativa en Tiempo 1 ni en Tiempo 3. En Tiempo 2, las aulas del grupo de intervención adhirieron más a las estrategias de manejo de grupo de UBC que las aulas del grupo de comparación y la diferencia fue estadísticamente significativa ($p < 0.01$). En términos de magnitud, el tamaño efecto del tratamiento en tiempo 2 fue de 0.59, lo que es considerado medio (Cohen, 1988) (ver Figura 4). Sin embargo, las aulas de tratamiento implementaron un 13% de todas las estrategias de alfabetización propuestas por el programa UBC, lo que representa una proporción relativamente pequeña.

Tabla 15

Análisis de regresión, diferencias en adherencia a estrategias de manejo de grupo UBC entre grupos de tratamiento según tiempo de medición

	Adherencia a manejo de grupo UBC Tiempo 1	Adherencia a manejo de grupo UBC Tiempo 2	Adherencia a manejo de grupo UBC Tiempo 3
Tratamiento	0.014 (0.013)	0.041** (0.016)	0.010 (0.015)
Intercepto	0.093*** (0.010)	0.088*** (0.012)	0.106*** (0.011)
N	90	87	94
CIC	0.000	0.000	0.164

Nota. Errores estándar en paréntesis. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ~ $p < 0.1$. CIC = Correlación Intra Clase.

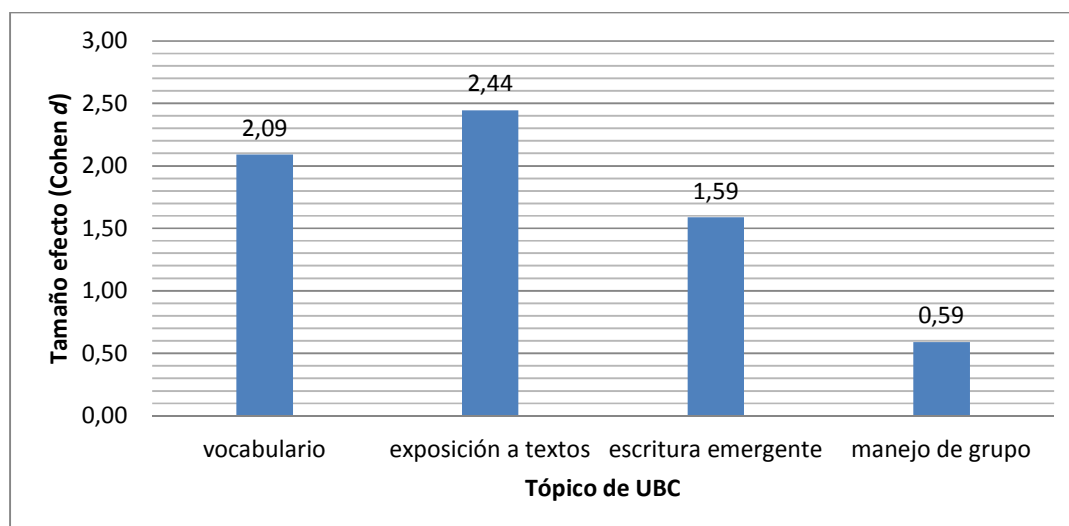


Figura 4. Tamaño Efecto del Tratamiento sobre Adherencia a Estrategias UBC en Tiempo 2.

Nota. El tamaño efecto de comprensión oral no pudo calcularse debido a que la desviación estándar del grupo control en tiempo 2 fue igual a 0.

Finalmente, respecto a la adherencia al uso de alcohol gel, la diferencia entre las aulas del grupo de tratamiento y control no fue estadísticamente significativa en tiempo 1 ni 2, pero sí lo fue en tiempo 3 a favor del grupo que recibió el programa ($p < 0.01$).

Resultados de modelos de crecimiento longitudinal. Como se explicó anteriormente este tipo de análisis puede hacerse solo con adherencia a estrategias de manejo de grupo, ya que no hubo suficiente poder estadístico para conducir este análisis con adherencia a estrategias de alfabetización UBC. Como se aprecia en la Tabla 16, la asignación al tratamiento incrementó el nivel promedio de exposición a los niños a estrategias de manejo de grupo UBC desde el inicio de prekínder al fin de kínder (Modelo 2, $\text{Tratamiento} = 0.023$, $p < .05$). Esto significa que la trayectoria de crecimiento de adherencia a manejo de grupo para el grupo de aulas que recibió el programa fue significativa.

Tabla 16.
Modelos de crecimiento longitudinal de adherencia a estrategias de manejo de grupo UBC (con efectos aleatorios para las escuelas)

	Modelo 1, Adherencia a manejo de grupo UBC: Solo tiempo	Modelo 2, Adherencia a manejo de grupo UBC: Modelo final
Tiempo	0.006 (0.004)	0.006 (0.004)
Tratamiento		0.023 [*] (0.011)
Intercepto	0.103 ^{***} (0.007)	0.090 ^{***} (0.009)
Nivel 2 Varianza Escuela	0.000 ^{***} (0.000)	0.000 ^{***} (0.000)
Nivel 2, Varianza en Tasa de cambio	0.000 ^{**} (0.000)	0.000 ^{**} (0.000)
Nivel 2, Varianza en status inicial	0.001 ^{***} (0.000)	0.001 ^{***} (0.000)
Nivel 1, Varianza residual de Aulas	0.003 ^{***} (0.000)	0.003 ^{***} (0.000)
N	271	271

Nota. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ~ $p < 0.1$. Errores estándar en paréntesis.

Pregunta de investigación 2 - ¿Las características de las educadoras (educación, edad, experiencia) predicen la dosis y adherencia de implementación de UBC en tiempo 2?

Resultados de cada tiempo de medición. Se exploró si la dosis a UBC varió según características de las educadoras en tiempo 1 y 2, y si el efecto de las características de las educadoras varió por condición de tratamiento. Debido a que la mayoría de las educadoras no se mantuvieron en la misma sala entre tiempo 2 y tiempo 3, no se pudo conducir este análisis en tiempo 3. Se controló por número de minutos codificados y se usaron interceptos aleatorios para escuelas, a fin de ajustar las estimaciones y errores estándar para los conglomerados de aulas al interior de las escuelas. Las características de las educadoras examinadas fueron edad de las educadoras, experiencia en aula de la educadora (en los niveles menos de 5 años,

5 a 14 años, y 14 años o más), si la educadora tuvo experiencia enseñando en una escuela particular pagada, y si la educadora tuvo formación de postítulo o diplomado.

No se encontró evidencia que la dosis de alfabetización UBC variara según características de las educadoras en tiempo 2, o que el efecto de las características de las educadoras en la dosis de alfabetización de UBC variara por condición de tratamiento.

Respecto a adherencia a estrategias de alfabetización, debido a la reducción de la muestra a las aulas que implementaron algún evento en los tópicos de alfabetización y la consecuente pérdida de poder estadístico, no se pudo conducir este análisis. Para adherencia a manejo de grupo, se encontró el mismo patrón de resultados que en dosis, sin relaciones significativas.

Resultados de modelos de crecimiento longitudinal. Debido a que el estudio UBC siguió a los niños en el tiempo y no a las educadoras, no fue posible conducir modelos de análisis de crecimiento longitudinal para este reporte. No hubo suficientes profesores que se mantuvieron el mismo tiempo entre tiempo 2 y 3 para hacer estos análisis posible (N= solo 37 educadoras o el 39% del total de educadoras en tiempo 3).

Pregunta de investigación 3.a - ¿El cambio en dosis de alfabetización UBC explica el crecimiento de los niños en lenguaje, controlando por características de los niños y de calidad de las oportunidades de aprendizaje ofrecidas por las educadoras?

Resultados de cada tiempo de medición. La tabla 17 resume los resultados para las relaciones entre dosis de tópicos de alfabetización UBC en Tiempo 2 y Tiempo 3, y resultados de desarrollo infantil. Los modelos controlaron por características sociodemográficas de los niños, profesionales de las educadoras, por la calidad del apoyo instruccional, y por los niveles de desarrollo obtenidos por los niños en el Tiempo 1. Los modelos también incluyeron interceptos aleatorios para las escuelas.

Se encontró que dosis de alfabetización de UBC no predijo resultados de lenguaje, alfabetización, ni Problemas Aplicados en Tiempo 2 (Tabla 17). Sin embargo, en Tiempo 3, mayor dosis de alfabetización de UBC estuvo positiva y estadísticamente asociada con resultados de Reconocimiento de Letras y Palabras, y de Dictado ($p < 0.01$) de los niños. Traduciendo estas asociaciones positivas a tamaños de efecto estandarizados (d de Cohen), el tamaño de efecto para Reconocimiento de Letras y Palabras fue mayor (de 0.20) que para Dictado (0.11), aunque en ambos casos se trató de un efecto pequeño.

Tabla 17.

Regresiones entre dosis de alfabetización de UBC y resultados de los niños en alfabetización y matemática, en Tiempo 2 y Tiempo 3.

	Vocabulario de imágenes	Identificación de Letras y Palabras	Dictado	Comprensión de textos	Problemas Aplicados
Dosis de UBC, T2	0.010 (0.011)	-0.013 (0.011)	-0.001 (0.006)	-0.002 (0.004)	-0.001 (0.012)
Dosis de UBC, T3	0.001 (0.012)	0.061** (0.023)	0.018** (0.006)	0.004 (0.007)	0.017 (0.014)

Nota. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ~ $p < 0.1$. Errores estándar en paréntesis.

Resultados de Modelos de crecimiento longitudinal. En modelos jerárquicos de crecimiento longitudinal que controlaron por características de los niños y de las educadoras y

que incluyeron efectos aleatorios para escuelas, se encontró que el incremento de un minuto en dosis de tópicos de alfabetización UBC estuvo asociada a un incremento de 0.01 puntos en la prueba de Dictado infantil ($p < 0.05$; d de Cohen = 0.07). El efecto de dosis UBC en los puntajes de Identificación de Letras y Palabras de los niños y en Problemas Aplicados variaron según el Tiempo ($p < 0.05$). Como se aprecia en las Figuras 5 y 6, niños expuestos a mayor dosis de tópicos de alfabetización de UBC (definida como una desviación estándar sobre el Promedio, es decir 11,37 minutos en tiempo 1; 20,24 en tiempo 2; y 21,72 en tiempo 3) tuvieron menores valores en promedio en Tiempo 1 para Identificación de Letras y Palabras y Problemas Aplicados, pero mayores valores en promedio en Tiempo 3 que los niños expuestos a nula dosis de tópicos de alfabetización UBC (definida como una dosis de cero minutos de UBC).

Si transformamos la contribución que dosis de alfabetización UBC ejerce sobre reconocimiento de letras y palabras a equivalencia de edad, los niños que estuvieron en aulas con nula dosis, tuvieron puntajes equivalentes a 4 años y 5 meses al inicio de prekinder, y a fin de kínder equivalente a 6 años, comparados con la muestra de referencia latinoamericana⁸. En cambio, los niños que estuvieron en aulas que mostraron alta dosis a lo largo del tiempo, partieron al inicio de prekinder con puntajes equivalentes a 4 años en la misma competencia, y lograron compensar la diferencia de 5 meses iniciales hasta alcanzar una edad equivalente a 6 años al fin de kínder.

No hubo relaciones significativas entre dosis de tópicos de alfabetización UBC y las otras medidas de los niños de lenguaje y alfabetización.

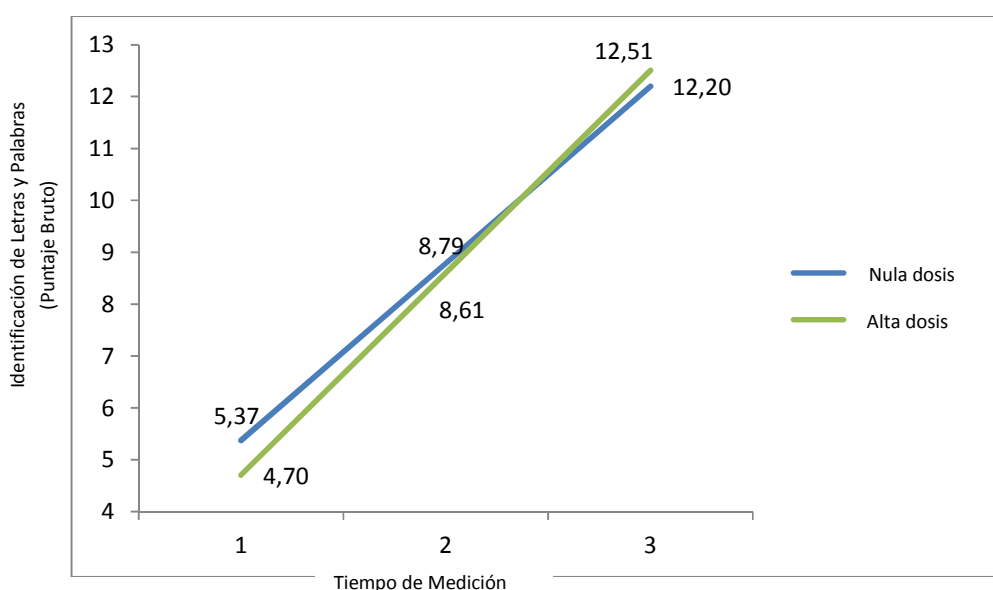


Figura 5. Efecto de Dosis de Alfabetización UBC a lo largo del Tiempo en el Puntaje de los Niños en Identificación de Letras y Palabras.

⁸ La batería Woodcock-Muñoz permite transformar los puntajes bruto promedio obtenidos, a edad de acuerdo a las normas obtenidas en una muestra latinoamericana de referencia. Por ejemplo, se puede decir que si la muestra obtuvo un promedio de 19 puntos bruto en la prueba de vocabulario, esta presenta un nivel de desarrollo equivalente a 4 años y 0 meses en vocabulario. Para realizar las transformaciones se utilizó el manual de la batería (Muñoz-Sandoval, Rueda, & Alvarado, 2005; Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005)

Transformando el efecto de dosis de alfabetización UBC sobre el crecimiento de los niños en problemas aplicados a equivalencia de edad, los niños que estuvieron en aulas con nula dosis, tuvieron puntajes equivalentes a 3 años y 5 meses al inicio de prekínder, y a fin de kínder puntajes equivalentes a 5 años y 4 meses, comparados con la muestra de referencia latinoamericana. En cambio, los niños que estuvieron en aulas con alta dosis a lo largo del tiempo, partieron al inicio de prekínder con puntajes equivalentes a 3 años y 2 meses en la misma competencia, y lograron compensar la diferencia de 3 meses iniciales hasta alcanzar también una edad equivalente a 5 años y 4 meses al fin de kínder.

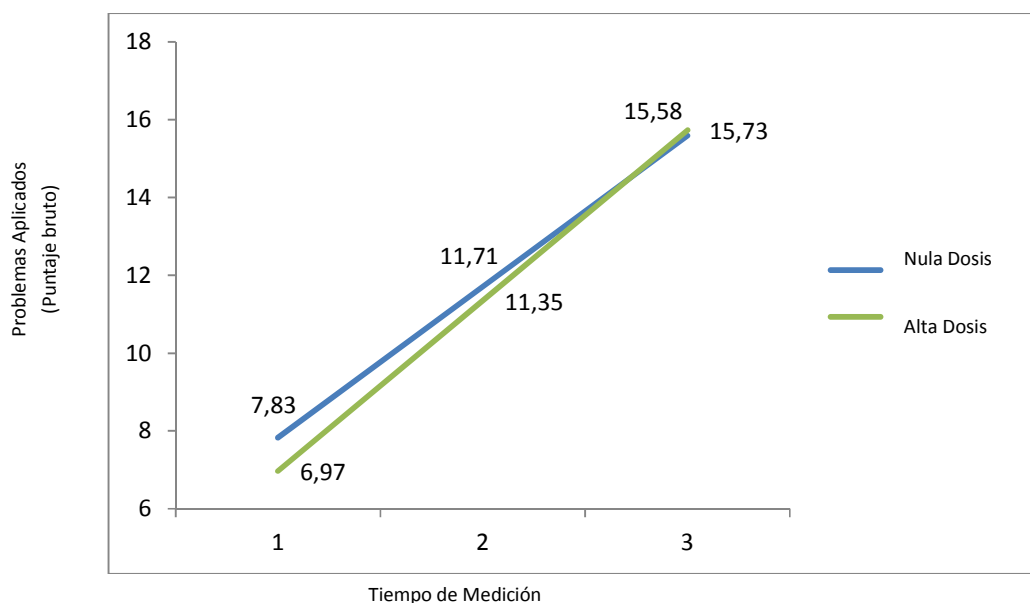


Figura 6. Efecto de Dosis de Alfabetización de UBC a lo largo del Tiempo en el Puntaje de los Niños en Problemas Aplicados.

Pregunta de investigación 3.b - ¿El cambio en adherencia a estrategias UBC de manejo de grupo, explican el crecimiento de los niños en variables socioemocionales?

La tabla 18 resume los resultados para la relación entre adherencia a manejo de grupo UBC en tiempo 2 y tiempo 3 y resultados socioemocionales de los niños. Los modelos controlaron por características de los niños y de las educadoras, por los puntajes de los niños obtenidos en tiempo 1. Asimismo, los modelos incluyeron interceptos aleatorios para las escuelas.

Adherencia a las estrategias de manejo de grupo no mostró efecto en los resultados de los niños, excepto por un efecto negativo encontrado en conducta positiva (medida por la observación del evaluador externo durante la situación e evaluación) en tiempo 3 ($p < 0.05$). Nuestra interpretación es que este efecto es espurio, dado el alto número de pruebas de significación estadísticas aplicadas, situación en la que se espera un resultado 'falso positivo' (Schochet, 2008). Previo a la publicación de estos resultados se espera probar esta hipótesis conduciendo nuevamente estos análisis con menos variables de desarrollo socioemocional que se obtendrían a partir de un análisis factorial confirmatorio.

Tabla 18.

Asociaciones entre resultados socio emocionales de los niños y adherencia a estrategias de manejo de grupo UBC en tiempo 2 y tiempo 3

	Control de Impulsos (evaluador)	Conducta Positiva (evaluador)	Comportamiento Externalizante (reporte de padres)	Comportamiento Prosocial (reporte de padres)
Adherencia UBC, T2	-0.201 (0.229)	-0.061 (0.219)	0.417 (0.364)	-0.226 (0.303)
Adherencia UBC, T3	-0.039 (0.277)	-0.725 (0.288)	0.150 (0.405)	0.025 (0.380)

Nota. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ~ $p < 0.1$. Errores estándar en paréntesis.

Resultados de modelos de crecimiento longitudinal. En modelos jerárquicos de crecimiento longitudinal que controlaron por características de los niños, de las educadoras, por calidad del apoyo instruccional en tiempo 1, y que incluyeron efectos aleatorios para escuelas, se encontró que el efecto de adherencia a estrategias de manejo de grupo UBC en control de impulsos de los niños varió a lo largo del tiempo ($p < 0.05$). Como se aprecia en la Figura 7, los niños que estuvieron en aulas expuestos a alta adherencia (definido como una desviación estándar sobre el promedio) tuvieron mayor control de impulsos en tiempo 1. En tiempo 3 los niños presentes en el grupo de aulas expuestos a alta adherencia a estrategias de manejo de grupo mostraron control de impulsos equivalentes a quienes estuvieron en el grupo de aulas con nula adherencia a estas estrategias. En términos prácticos, pese a la diferencia significativa en tiempo 1, la diferencia entre mayor y menor adherencia a estrategias de manejo de grupo a lo largo del tiempo no parece ejercer un efecto sobre el control de impulsos de los niños.

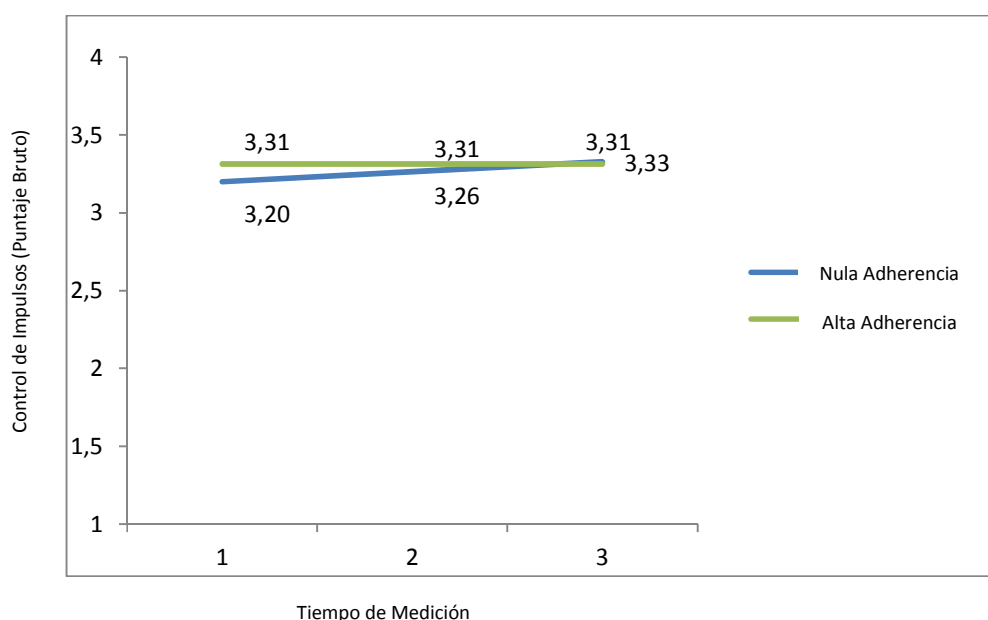


Figura 7. Efecto de Adherencia de Manejo de Grupo a UBC a lo largo del Tiempo en Resultados de los Niños en Control de Impulsos.

Conclusiones

Este estudio exploró si durante el período de dos años del estudio de UBC, las aulas tanto del grupo que recibieron la intervención, como las que no, han aumentado tanto el tiempo que destinan a trabajar los tópicos de alfabetización del programa (vocabulario, comprensión oral, escritura emergente, exposición a textos, y conocimiento de lo impreso), es decir la *dosis de alfabetización UBC*, como la *adherencia* a las estrategias de alfabetización y manejo de grupo que el programa promovió. Asimismo, exploró si las diferencias en la dosis y adherencia de las aulas de ambas condiciones de tratamiento pueden atribuirse a características de desarrollo profesional de las educadoras, y si el cambio de la dosis y adherencia mostró una asociación con el crecimiento en los resultados de los niños.

Gracias al diseño experimental de este estudio, se puede concluir en primer lugar que el programa UBC mostró efectividad para modificar prácticas de educadoras particularmente durante prekinder (NT1) en términos de dosis y adherencia. Las educadoras que recibieron el programa comparadas con las que no, aumentaron la dosis de tópicos de alfabetización relevantes, y adhirieron notoriamente más a la modalidades propuestas por UBC para implementar dichos tópicos por un lado, y por otro, para manejar la conducta de los niños a favor de la generación de un ambiente propicio para el aprendizaje. En tiempo 2, los 5.2 minutos más observados en el grupo que recibió el programa, representan en una jornada diaria alrededor de 12.1 minutos más de dosis de alfabetización de UBC. En Tiempo 3, el número equivalente a una jornada diaria fue de 8.96 minutos más de dosis de alfabetización de UBC en las salas del grupo de intervención, en relación a las del grupo de comparación.

De acuerdo al concepto *procesos próximos* del modelo bioecológico del desarrollo humano (Bronfenbrenner & Morris, 2006), se requiere que procesos de enseñanza ocurran con cierta extensión y reiteración en el tiempo para impactar en los resultados de desarrollo y aprendizaje de los niños. En efecto, en este estudio se encontró que aumentar el tiempo de exposición a los niños a tópicos de alfabetización de UBC, se relaciona con el crecimiento de las capacidades de escritura emergente (prueba de Dictado), y reconocimiento de letras y palabras. Especialmente una exposición aproximada de 20 minutos en prekinder y kínder en los tópicos de alfabetización UBC, aceleró la trayectoria de crecimiento de los niños en la capacidad de reconocer letras y palabras a fines de kínder, compensando incluso la diferencia de edad de 5 meses que los niños del grupo de alta dosis tenían a inicios de prekinder, en comparación con los niños que por el mismo período de tiempo no fueron expuestos a dosis en estos tópicos de alfabetización.

Más aún, se observó un efecto similar del aumento de dosis de alfabetización UBC en la aceleración del crecimiento en la capacidad de problemas aplicados. La relación positiva entre alta dosis de alfabetización UBC y la posterior competencia de problemas aplicados infantil, no debe interpretarse como un resultado sorpresivo. Tanto teóricamente, como evidencia reciente sugieren que mayor desarrollo en habilidades de lenguaje y alfabetización inicial pueden apoyar el aprendizaje de habilidades matemáticas de los niños, y vice versa (Duncan et al., 2007; Harrison, McLeod, Berthelsen, & Walker, 2009; Wagner, Venezky & Street, 1999).

Pese a la evidencia de efectividad del programa UBC para incrementar la dosis de alfabetización inicial, no debe perderse de vista que en el segmento escolar estudiado, aun estamos en un escenario de escaso tiempo instruccional destinado a desarrollar competencias de alfabetización inicial fundamentales. De un total de 95 minutos observados en promedio, las aulas del grupo de intervención implementaron tópicos de alfabetización inicial durante aproximadamente 12 minutos a fines de prekinder y kínder, y las salas del grupo de comparación expusieron a sus niños a solo 8 y 9 minutos a fines de prekinder y kínder respectivamente. Comparado con evidencia internacional, esta extensión de tiempo difícilmente logrará impactar positivamente competencias de alfabetización inicial de los niños.

Por ejemplo, el proyecto Early Reading First ha propuesto comprometer a educadoras y directivos de las escuelas con 2,5 horas diarias destinadas a instrucción de lenguaje y alfabetización inicial en el nivel parvulario (Abbott, Atwater, Lee, & Edwards, 2011). En coherencia con el programa Early Reading First, en aulas preescolares de escuelas públicas en Boston, se implementó el currículo 'Opening the World of Learning' (OWL) (Schickedanz & Dickinson, 2005, en Weiland, 2012) que estipulaba 3 horas diarias de instrucción en lenguaje y alfabetización, y especificaba cuanto tiempo los niños debían pasar en actividades como círculo, tiempo en rincones, relato de historias y además especificaba qué debería ocurrir en cada momento. De acuerdo al juicio de los evaluadores de la implementación del programa, el currículum fue implementado en un nivel moderado a alto, y se encontró impactos sustanciales en habilidades de lenguaje y alfabetización (Weiland, 2012). Si bien estos estudios pueden considerarse una referencia, aún se requiere más investigación para determinar una prescripción basada en evidencia, de la dosis necesaria de trabajo en alfabetización inicial para impactar en resultados de los niños chilenos. Para determinarlo se requeriría conducir estudios experimentales (randomized controlled trial) que asignen aleatoriamente grupo de aulas a distintas dosis de exposición a estos tópicos y comparar las diferencias de los distintos grupos de tratamiento en los resultados de desarrollo de los niños.

Como recomendaciones de política pública a partir de los hallazgos de dosis de este estudio, se sugiere a nivel de sistema educativo que exista una expectativa explícita de dosis para trabajar al menos los tópicos de alfabetización inicial. La propuesta de política en el nivel curricular también ha sido sugerida internacionalmente para el nivel preescolar: "secuencias de aprendizaje altamente estructuradas pueden favorecer un manejo de grupo efectivo y (...) asegurar que los niños pasen relativamente más tiempo en actividades de alfabetización inicial" (Phillips, Gormley, & Lowenstein, 2007, p.26). Más específicamente, se sugiere que en instrumentos curriculares como los 'Programas Pedagógicos' se comunique explícitamente un tiempo mínimo esperable para trabajar los diversos aprendizajes esperados. Esto a la fecha no está presente ni en los programas pedagógicos, ni en ninguno de los elementos de diagnóstico o seguimiento de los planes de mejoramiento de la ley SEP (cuyas escuelas participantes mantienen un perfil socioeconómico similar a las participantes en este estudio). La omisión de este elemento a nivel de sistema, se retroalimenta por la posición 'anexa' que se ha diagnosticado ocupa el nivel parvulario en el contexto de las escuelas en Chile, por la falta de dominio, involucramiento y monitoreo pedagógico que mantienen los jefes de UTP en materias de educación parvularia, y las escasas iniciativas para que las educadoras de párvulos participen en instancias de reflexión o asesoría técnica (Facultad de Educación, UDP, 2010b). De ahí que señales claras de dosis desde el sistema educativo podría contribuir a compensar esta realidad de las escuelas que tardará en cambiar en forma global. La baja intensidad de exposición a experiencias de aprendizaje que promuevan competencias de alfabetización inicial fundamentales, documentada en los últimos 10 años y que no ha cambiado sustancialmente a juzgar por los resultados de dosis de alfabetización en tiempo 1, tiene efectos aún más preocupantes en los niños provenientes de contextos de mayor vulnerabilidad, quienes suelen recibir menor estimulación lingüística en el hogar (Hoff, 2006), y por lo tanto, son quienes sufren desde temprano y en forma acumulativa la inequidad en las oportunidades de aprendizaje (Dickinson, 2011).

Es importante además, que tomar conciencia sobre la variable dosis sea recibido y asumido en el currículo y evaluación de los programas formativos de las instituciones de formación inicial de educadoras de párvulos, a fin que desde una fase temprana de formación, las educadoras reconozcan la importancia de esta materia. El estado puede recurrir a figuras como convenios de desempeño con las universidades a fin de asegurar que estos elementos sean incorporados.

Para especificar la dosis esperada, es recomendable que desde el nivel central se promuevan estudios que entreguen evidencia respecto a dosis necesarias de enseñanza para

impactar en el aprendizaje de competencias fundamentales en el nivel preescolar, o que instituciones sin fines de lucro interesadas en contribuir a la educación parvularia financien estudios como los recomendados.

Para quienes contribuyen a la educación parvularia desde el ámbito de la investigación, instituciones privadas, sostenedores o de las mismas escuelas, si el foco es diseñar y evaluar efectos de programas de desarrollo profesional docente, se recomienda que desde el diseño del programa, se comunique una dosis esperable de implementación y además se establezcan mecanismos específicos para monitorear y retroalimentar este aspecto. Una alternativa sustentable es involucrar a los equipos directivos (Abbott, et al., 2011) en roles de monitoreo, aseguramiento de la implementación y entrega de retroalimentación permanente, o bien, en el rol de formación y apoyo de comunidades de aprendizaje profesional en las escuelas donde se planifique, observe y retroalimente mutuamente la dosis, adherencia y calidad de la implementación de tópicos de alfabetización inicial (Darling-Hammond, 2012).

Respecto a la adherencia al programa UBC, considerando los niveles de significación estadística y los tamaño efecto, se encontró que a fines de prekinder las aulas participantes del programa adhirieron en mayor medida a las estrategias propuestas por UBC para las dimensiones vocabulario, exposición a textos, comprensión oral, escritura emergente y manejo de grupo, en comparación con las aulas que no recibieron el programa. Responder esta pregunta es relevante pues también podría haberse encontrado alta adherencia del grupo que no recibió el programa, conocido como fenómeno de ‘contaminación del grupo de comparación’, o baja adherencia de quienes participaron en el programa, lo que se entiende como ‘baja receptividad’ a un programa (Durlak, 2010).

Es posible que la conjunción de contaminación del grupo de comparación (expresado en el aumento de los promedios de adherencia en tiempo 3 comparado con tiempo 2) y una disminución en la implementación de estrategias del programa en el grupo tratamiento (por el descenso de los promedios entre tiempo 3 y tiempo 2), explique la falta de diferencias significativas encontradas entre los grupos a fines de kínder (tiempo 3), al menos en adherencia a exposición a textos y comprensión oral. La contaminación en las aulas del grupo control en exposición a textos y comprensión oral puede tener origen en diversas iniciativas ministeriales impulsadas en años coincidentes con el período de implementación UBC.

En el año 2009 (período que coincide con el segundo año de intervención de la cohorte 1 de UBC) se comienzan a implementar los programas pedagógicos de Educación Parvularia para NT1 y NT2 (prekinder y kínder respectivamente) y se realizan capacitaciones a nivel nacional para incorporarlos en el trabajo cotidiano. Al incorporar ejemplos de experiencias de aprendizaje, este instrumento entregó lineamientos más claros que los presentes en las Bases Curriculares del año 2001, para que las educadoras expusieran a los niños a literatura infantil, favorecieran la comprensión de información explícita evidente de historias escuchadas en NT1 (MINEDUC, 2008a) y favorecieran la comprensión a través de predicciones y de un nivel inferencial en NT2 (MINEDUC, 2008b). La mayor claridad y especificación de los programas ha sido positivamente recibida por las educadoras, de acuerdo a un estudio de implementación de estos programas (Facultad de Educación UDP, 2010b). Se puede apreciar que los elementos mencionados de exposición a literatura infantil y de mediación de la comprensión oral están también presentes en la propuesta de UBC, aunque en UBC hay un mayor énfasis en el nivel de comprensión inferencial y en el uso de estrategias de comprensión. Estos elementos presentes en los Programas Pedagógicos, también se encuentran en los cuadernillos pedagógicos de la Estrategia LEM de lenguaje, existentes desde el año 2006, que solo tiene especificaciones para NT2, nivel en el que se ha propuesto que podría haber cierta influencia de programas nacionales que atenuaron las diferencias entre los grupos del presente estudio.

En el año 2008 comienza a regir la Ley de Subvención Escolar Preferencial (SEP) que asigna recursos adicionales por cada alumno prioritario y por concentración de los mismos en las escuelas. Esta asignación se condiciona a resultados educativos de aprendizaje que las

escuelas deben dar cuenta y al cumplimiento de procesos a los que la comunidad educativa se compromete para alcanzarlos (fuente: anexo 1 de www.planesdemejoramiento.cl). Entre los mecanismos propuestos por la ley, el MINEDUC propone orientaciones para que las escuelas puedan evaluar específicamente la comprensión lectora desde el nivel parvulario, reforzándose la necesidad de recurrir a los programas pedagógicos y de evaluar estas competencias (MINEDUC s/f). Las escuelas de este estudio tuvieron entre sus criterios de inclusión la presencia de un número importante de alumnos en condición de vulnerabilidad, por lo que se puede asumir que la influencia de estos mecanismos ministeriales operaron al menos en la mayor parte de las escuelas estudiadas, pudiendo disminuir nuevamente las diferencias en los tópicos mencionados en los grupos tratamiento y comparación. También es posible que el Plan de Fomento Lector que comenzó en el año 2011, haya reforzado aún más la tendencia ya descrita, aunque su efecto cubriría solo el último año de intervención de la cohorte 3. Finalmente se puede desestimar la eventual influencia del programa Plan de Apoyo Compartido, ya que de acuerdo a bases de datos provistas por el departamento de estadísticas del MINEDUC, solo 3 salas con un total de 64 niños del grupo de comparación de este estudio participaron en PAC durante el año 2011.

Un elemento adicional que puede explicar el alza de dosis de tópicos de alfabetización UBC y de adherencia en comprensión oral y exposición a textos en tiempo 3, es que en el contexto de funcionamiento de la ley SEP, que exige demostración de resultados a través de mejoras en los puntajes SIMCE en 4° básico, los equipos directivos ejerzan mayor demanda de implementación curricular en kínder que en prekínder, que desde una expectativa de rendición de cuentas, es el año que debiera prepararse a los niños para que en el primer año básico puedan 'aprender a leer'. En cambio, en un posible escenario de menor presión sobre el prekínder, es posible que las aulas del grupo de comparación tiendan a implementar el currículo en menor medida, y la efectividad del modelo de capacitación y acompañamiento del programa UBC haya marcado más claramente la diferencia en el grupo de aulas que recibieron el programa.

La leve disminución observada en adherencia a estrategias UBC en los tópicos vocabulario, comprensión oral y exposición a textos a fines de kínder, es un fenómeno reportado frecuentemente en los estudios de seguimiento de efectos de tratamiento. Esto pone en evidencia que los programas que intervienen en forma externa a la escuela pueden ser efectivos, pero dicha efectividad puede ser aún más sustentable en el tiempo si parte del modelo de cambio incluye roles de monitoreo y apoyo dentro de las mismas comunidades educativas (Abbott, et. al., 2011; Darling-Hammond, 2012).

La explicación propuesta para comprender el alza de puntaje de adherencia en el grupo de comparación en comprensión oral y exposición a textos, también sirve para explicar el hecho que entre las educadoras del grupo de tratamiento la mayor adherencia observada sea en comprensión oral. También debe considerarse que en la distribución de contenidos del programa UBC, comprensión oral fue cubierto por más tiempo, ocupando dos de 6 ciclos de capacitación cada año, a diferencia de un ciclo para vocabulario y uno de escritura emergente al año. Por otra parte, la mayor dosis y adherencia a estrategias de vocabulario a fines de prekínder es un resultado que merece destacarse dado en primer lugar, la relevancia crucial que el desarrollo de vocabulario temprano tiene para el desempeño posterior en comprensión de textos (NICHD Early Child Care Research Network, 2005), como para la comprensión de otros contenidos escolares y rendimiento académico posterior (Storch, & Whitehurst, 2002); y en segundo lugar, dada los hallazgos recientes de escaso tiempo destinado a la instrucción explícita de vocabulario en aulas subvencionadas del nivel parvulario en Chile (Eyzaguirre & Fontaine, 2008; Facultad de Educación PUC, 2011; Strasser, Lissi, & Silva, 2009), y la necesidad expresada por los equipos pedagógicos del nivel parvulario, de conocer estrategias concretas que estimulen las competencias de vocabulario, que según lo que reportan no reciben de su formación inicial (Facultad de Educación PUC, 2011).

Características de la propuesta pedagógica de vocabulario que ofrece el programa UBC merecen ser analizadas con mayor detalle, pues se diferencia en forma importante de las sugerencias generales presentes en los Programas Pedagógicos (MINEDUC, 2008a, 2008b), de la propuesta LEM lenguaje de NT2 (MINEDUC, 2006), del programa LEAMOS (Facultad de Educación UDP, 2010a), o en la escasa alusión a vocabulario en lineamientos de SEP (MINEDUC, s/f.), y de acuerdo a lo reportado por Facultad de Educación, PUC (2011), es presumible que tampoco esté suficientemente incorporado en los currículos actuales de la formación de educadoras de párvulos en nuestro país. La propuesta opta por la enseñanza de un tipo de palabras particular, denominadas de nivel 2, que aparecen con frecuencia en el lenguaje oral y textos escritos, por lo tanto son relevantes para desarrollar la capacidad de comprensión oral de textos literarios de calidad. El significado de las palabras de nivel 2 no son de fácil inferencia por los niños pequeños, y pueden identificarse porque puede encontrarse una palabra que represente el mismo significado en el lenguaje cotidiano (Beck, McKeown, & Kucan, 2005). Por ejemplo, los niños pueden enriquecer su léxico respecto al campo de contextos en que pueden utilizar la palabra de uso frecuente 'mirar' (o palabra de nivel 1), al aprender la palabra de nivel 2 'divisar', que le serviría para comprender que el cocodrilo '*se deslizó por la jungla hasta que divisó una montaña solitaria*' en la historia 'Los cocodrilos copiones' (Bolam, & Bedford, 2006) y para más tarde incorporar la palabra 'divisar' en su habla cotidiana. Otros elementos relevantes de la propuesta es que la enseñanza de las palabras proviene de un texto leído previamente, que se planifica previamente la mejor manera de explicar el significado de una palabra compleja en forma accesible a la comprensión de los niños pero sin distorsionar su significado; que la educadora también planifica ejemplos en los que se puede aplicar la palabra, procurando aplicarla en diversos contextos; y finalmente también se planifica un momento en que los niños tengan oportunidad de interactuar con el significado de la palabra nueva, por ejemplo, invitar a los niños a salir al patio con unos binoculares estimulándolos a decir que es lo que divisan; lo que se hace necesario para garantizar la comprensión profunda del significado (Beck, McKeown, & Kucan, 2002). La propuesta implementada por el programa ha reportado evidencia de beneficio en el aprendizaje de palabras de niños en contextos de vulnerabilidad (Beck, & McKeown, 2007).

Pese a la complejidad y novedad de la propuesta descrita, la diferencia significativa en adherencia a las estrategias de vocabulario descritas a fines de prekindergarten a favor del grupo de aulas que participó en UBC, muestra que es una propuesta viable de ser aprendida y luego implementada por iniciativa propia de las educadoras que trabajan en escuelas municipales con alto número de niños en situación de vulnerabilidad. Este dato es aún más destacable si consideramos que el nivel de calidad instruccional, de acuerdo a la medición CLASS, que mostraron tener las educadoras participantes del estudio fue particularmente bajo (Yoshikawa et al., 2012), lo que indica que dichas características de entrada no constituirían un obstáculo.

De los resultados presentados se propone como segunda recomendación de política pública, la necesidad de generar una política nacional que ponga prioridad en el desarrollo del vocabulario en el nivel parvulario. Una política debiera identificar diversos niveles de actuación, entre ellos, que desde el nivel de sistema educativo se incorporen propuestas más específicas para trabajar vocabulario en los Programas Pedagógicos y que el aprendizaje de palabras de nivel 2 sea un aprendizaje esperado para ambos niveles de NT. Para el nivel de actores que contribuyen a la educación parvularia desde el diseño y evaluación de programas de intervención (investigadores, entidades sin fines de lucro), se recomienda diseñar, implementar y evaluar el impacto de intervenciones acotadas únicamente al entrenamiento de estrategias de vocabulario como las propuestas por UBC, y también con una clara prescripción de dosis de implementación en aula con los niños, para probar en qué medida una intervención como esta y de una intensidad mayor que la probada con UBC impacta en los resultados de los niños.

Para el nivel de formación inicial de educadoras de párvulos, se recomienda que existan cursos mínimos de carácter teórico y práctico que aseguren el desarrollo de competencias para

que las futuras educadoras realmente implementen trabajo pedagógico en vocabulario; que conozcan criterios para seleccionar palabras de nivel 2 y manejen diversas estrategias para asegurar la comprensión del significado en los niños. Intervenir en este nivel es crucial, dados el bajo nivel de conocimiento de didáctica específica en materia de alfabetización inicial (Facultad de Educación UDP, 2010 a), y la baja calidad de las oportunidades de aprendizaje que se imparte en el nivel parvulario (Yoshikawa, et al., 2012), de acuerdo a ello se recomienda que el MINEDUC considere la movilización de este cambio a través del programa de Convenios de Desempeño para impactar en el nivel de educación superior.

En este estudio se encontró ausencia de relación entre adherencia a estrategias de manejo de grupo UBC y resultados de desarrollo socioemocional de los niños. Sin embargo, este hallazgo debe considerarse preliminar y aun no concluyente respecto a la posible contribución de adherencia a las estrategias de manejo de grupo de UBC en el desarrollo socioemocional de los niños. Es conveniente realizar a futuro análisis factorial de todas las medidas de desarrollo socioemocional a fin de reducir las a factores o constructos latentes que permitan por un lado, remover el error de las variables manifiestas (Kline, 2004) utilizadas en este estudio, y por otro, reducir la cantidad de pruebas de significación (Schochet, 2008); y luego, explorar nuevamente la relación entre adherencia y los factores de desarrollo socioemocional.

Las características profesionales de las educadoras utilizadas en este estudio no estuvieron asociadas con los niveles de dosis ni de adherencia de alfabetización UBC a la que expusieron a sus niños. Este hallazgo indica que un programa de desarrollo profesional docente como el estudiado tiene el potencial de adaptarse a la diversidad de las educadoras de este estudio tanto en términos de edad, años de experiencia en aula, participación en experiencias de postítulo o diplomado, y experiencia previa en establecimientos de dependencia privada. De todos modos, es conveniente continuar explorando si otras variables de la trayectoria profesional pueden predecir la adherencia y dosis de las educadoras al programa. Asimismo, es posible que otras variables más vinculadas a procesos cumplan un rol en resultados diferenciales de adherencia o dosis. En futuros análisis que consideren solo el grupo que recibió el programa, se podrá explorar por ejemplo, si la facilitadora que acompañó a las educadoras predijo el nivel de dosis y adherencia.

Si bien se ha postulado que la observación es la evidencia más directa para capturar la práctica pedagógica y que ha ofrecido mayor poder para explicar la varianza de resultados de los niños (Hamre et al., 2010), se podría considerar una limitación de este estudio cierto sesgo de deseabilidad social en los hallazgos, ya que las educadoras conocían el día en que serían filmadas. De acuerdo a ello, se podría hipotetizar que los resultados obtenidos muestran el mejor desempeño que una educadora puede demostrar en una situación de evaluación (Danielson, 1996), en vez que una aproximación válida a la dosis y adherencia cotidiana de UBC a la que exponen a los niños con quienes trabajan. Esta hipótesis recibiría mayor apoyo si no hubiésemos encontrado una relación entre dosis y resultados de los niños, tanto en los resultados de corte transversal como de crecimiento longitudinal. Sin embargo, ese no fue el caso. Por lo tanto, se puede considerar que la medida de dosis capturada en este estudio es una adecuada aproximación a la dosis de alfabetización inicial a la que cotidianamente las educadoras expusieron a sus niños. Desafortunadamente no pudo explorarse la relación entre adherencia a UBC y el crecimiento en los resultados de los niños, por lo que no puede descartarse la influencia de deseabilidad social en la medida de adherencia a las estrategias UBC, constituyendo una limitación de este estudio.

Una segunda limitación de estudio es que al haber aprovechado un diseño de investigación previo, no fue posible controlar la permanencia de las educadoras a lo largo de los tres momentos de medición, lo que habría permitido responder si alguna de las características de las educadoras predecía el cambio en adherencia y dosis. Recientemente se ha planteado que esta limitación está presente en general en la literatura sobre los efectos de programas de desarrollo profesional docente en el nivel parvulario (Hamre, Downer, Jamil, & Pianta, 2012),

por lo que se recomienda que en el diseño de futuros estudios con preguntas de investigación similares se considere el seguimiento no solo de los niños sino también de las educadoras.

Este estudio ha contribuido con evidencia robusta a la comprensión de los cambios en prácticas de educadoras de párvulos en experiencias de desarrollo profesional docente en Chile, y ha entregado nueva evidencia sobre el uso del tiempo en tópicos de alfabetización relevantes, generalizable a escuelas de nivel municipal que atienden a niños en situación de vulnerabilidad.

Referencias

- Abbott, M., Atwater, J., Lee, Y., & Edwards, L. (2011). A Data-Driven Preschool PD Model for Literacy and Oral Language Instruction. *NHSA Dialog*, 14(4), 229-245.
- Alvarado, C.G., Ruef, M.L., & Schrank, F.A. (2005). *Comprehensive Manual: Woodcock-Munoz Language Survey-Revised*. Itasca, IL: Riverside Publishing.
- Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge, UK: Cambridge University Press.
- Beck, I. L. & McKeown, M.G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, 107(3), 251-271.
- Beck, I.L, McKeown, M.G. & Kucan, L. (2002). *Bringing Words to life. Robust Vocabulary Instruction*. New York: Guilford Press.
- Beck, I. L., McKeown, M. G. & Kucan, L. (2005). Choosing words to teach. In E. H. Hiebert & M. L. Kamil (Eds.). *Teaching and learning vocabulary: Bringing research to practice*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Bellei, C, Raczynski, D., & Osses, A. (2010). ¿Qué hemos aprendido sobre programas de asistencia técnica educativa? En Bellei, C. (Ed.) *Asistencia Educativa en Chile: ¿Aporte al mejoramiento Escolar?* (pp. 31-76). Santiago de Chile: Editorial Ocho libros.
- Bierman, K.L., Domitrovich, C.E., Nix, R.L., Gest, S.D., Welsh, J.A., Greenberg, M.T., Blair, C., Nelson, K. & Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI Program. *Child Development*, 79(6), 1802 – 1817.
- Bolam, E. & Bedford, D. (2006). *Los cocodrilos copiones*. Caracas: Ekaré.
- Bronfenbrenner, U., & Morris, P.A. (2006). The bioecological model of human development. In R. Lerner, & W. Damon (Eds.), *Handbook of Child Psychology (6th ed.): Vol-1, Theoretical models of human development*. (pp. 793-828). Hoboken, NJ: John Wiley & Sons.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032-1053.
- Century, J., Rudnick, M., & Freeman. C. A (2010) Framework for Measuring Fidelity of Implementation: A Foundation for Shared Language and Accumulation of Knowledge *American Journal of Evaluation*, 31, 199-218.
- Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal*, 45, 443-494.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.
- Conduct Problems Prevention Research Group (CPPRG). (1990). *Social Competence Scale (Teacher Version)*. Available from the Fast Track Project Web site, <http://www.fasttrackproject.org>
- Connor, C. M., Morrison, F. J. & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, 98, 665-689.

- Cooper, D.H., Roth, F.P. & Speece, D.L. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research, 95*, 259-272.
- Cordray, D.S. & Hulleman, C. (2009). Implementation fidelity and Achieved Relative Strength in Randomized Controlled Trials. Invited presentation for the IES 4th Annual Research Conference. Washington, DC, June 4, 2009.
- Cordray, D.S. & Pion, G.M. (2006). Treatment strength and integrity: Models and methods. In R. Bootzin & P. McKnight (Eds.), *Strengthening Research Methodology: Psychological Measurement and Evaluation*. (pp. 103-124) Washington, D.C.: American Psychological Association.
- Danielson, C. (1996). *Enhancing Professional Practice: A Framework for teaching* (1st ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (2012). Desarrollo de un enfoque sistémico para evaluar la docencia y fomentar una enseñanza eficaz. *Pensamiento Educativo, 49*, 1-20.
- Darrow, C. & Dickinson, D. (2011). The Development and Application of Fidelity Measures in a Preschool Curriculum Intervención. Paper presented at the Fall Conference of the Society for Research in Educational Effectiveness, Washington, DC.
- Denham, S., Bouril, B., & Belouad, F. (1994). Preschoolers' affect and cognition about challenging peer situations. *Child Study Journal, 24*, 1 – 21.
- Diamond, A., Carlson, S. M., & Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology, 28*, 689-729.
- Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do". *Developmental Psychobiology, 29*(4), 315-334.
- Dickinson D. (2011). Teachers' Language Practices and Academic Outcomes of Preschool Children. *Science, 333*, 964-967. DOI:10.1126/science.1204526.
- Domitrovich, C., Gest, S., Jones, D., & Gill, S., & Sanford DeRousie, R. (2010). Implementation quality: Lessons learned in the context of the head start REDI trial. *Early Childhood Research Quarterly, 25*, 284-298.
- Domitrovich, C., Scott G., Sukhdeep, G., Damon J., & DeRousie, R. (2009). Individual Factors Associated With Professional Development Training Outcomes of the Head Start REDI Program. *Early Education & Development, 20*, 402 - 430.
- Downer, J., Locasale-Crouch, J., Hamre, B. & Pianta, R. (2009). Teacher Characteristics Associated With Responsiveness and Exposure to Consultation and Online Professional Development Resources. *Early Education & Development, 20*, 431 – 455.
- Duncan, G.J., Claessens, A., Huston, A.C., Pagani, L.S., Engel, M., Sexton, H., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428- 1446.
- Durlak, J. (2010). The Importance on Doing well in Whatever you do: A Commentary on the Special Section "Implementation Research in Early Childhood Education". *Early Childhood Research Quarterly, 25* (3), 348-357.

- Dusenbury L., Brannigan R., Falco M., & Hansen, W. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research* 18, 237-56.
- Eidelman, H., Weiland, C. & Yoshikawa, H. (2011). Fidelity-to-curriculum in a universal public preschool program: Psychometrics and links to children's outcomes. Poster presented at the biennial meeting of the Society for Research in Child Development, Montreal.
- Eyzaguirre, B. & Fontaine, L. (2008). *Las escuelas que tenemos*. Santiago, Chile: Centro de Estudios Públicos.
- Facultad de Educación PUC. (2011). Informe final: Alfabetización en establecimientos chilenos subvencionados. Santiago de Chile. Recuperado de <http://ded.mineduc.cl/mineduc/ded/documentos/Alfabetizacion%20-%20Informe%20final.pdf> el 24 de Noviembre de 2012.
- Facultad de Educación UDP (2010a). Evaluación de impacto del "Proyecto Leamos" Primer y Segundo Nivel de Transición de escuelas municipales y subvencionadas. Santiago de Chile. Recuperado de <http://www.educacion2.udp.cl/seminarios/201109/ppt/2010-ProyectoLeamos.pdf> el 21 de Enero de 2013.
- Facultad de Educación UDP (2010b). Seguimiento y evaluación de la implementación de los Programas Pedagógicos para el primer y segundo nivel de transición en escuelas municipales del país. Santiago de Chile. Recuperado de http://w3app.mineduc.cl/mineduc/ded/documentos/Evaluacion_Implementacion_Programas_Pedagogicos_NT1_NT2.pdf el 21 de Enero de 2013.
- Farran, D., Lipsey, M., Clements, D., Sarama, J., Hofer, K., Bilbrey, C., & Vorhaus, E. (2011). The Mechanisms behind the Results: Moderators of Building Blocks Curricular Effects. Paper presented at the Fall Conference of the Society for Research in Educational Effectiveness, Washington, DC.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10(4), 483-527.
- Goodson, Layzer, Smith & Rimdzius, (2004). The Observation Measures for Language and Literacy OMLIT. Non published document.
- Greenwood, C, Tapia, Y., Abbott, M. & Walton, C. (2003). A building-based case study of evidence-based literacy practices: Implementation, reading behavior, and growth in reading fluency, K-4. *Journal of Special Education*, 37(2), 95-110.
- Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Hamre, B. K., Downer, J. T., Jamil, F. M., & Pianta, R. C. (2012). Enhancing teachers' intentional use of effective interactions with children. In R. C. Pianta, W. S. Barnett, L. M. Justice & S. M. Sheridan (Eds.), *Handbook of early childhood education* (pp. 507-532). New York: Guilford Press.
- Hamre, B. K., et al. (2010). Implementation fidelity of MyTeachingPartner literacy and language activities: Association with preschoolers' language and literacy growth. *Early Childhood Research Quarterly*, 25 (3), 329-347.

- Harrison, L. J., McLeod, S., Berthelsen, D., & Walker, S. (2009). Literacy, numeracy, and learning in school-aged children identified as having speech and language impairment in early childhood. *International Journal of Speech-Language Pathology*, *11*, 392-403.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*, 55-88.
- Janus, M., & Offord, D., (2007). Psychometric properties of the Early Development Instrument (EDI): A teacher-completed measure of children's readiness to learn at school entry. *Canadian Journal of Behavioural Science*, *39*(1), 1-22.
- Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in pre-kindergarten programs serving at-risk pupils. *Early Childhood Research Quarterly*, *23*(1), 51-68.
- Justice, L., Mashburn, A., Pence, K., & Wiggins, A. (2008). Experimental Evaluation of a Preschool Language Curriculum: Influence on Children's Expressive Language Skills. *Journal of Speech, Language, and Hearing Research*, *51*, 983-1001.
- Kerr, D. M., Kent, L., & Lam, T. C. M. (1985). Measuring program implementation with a classroom observation instrument: The interactive teaching map. *Evaluation Review*, *9*, 461-482.
- Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development*, *67*(2), 490-507.
- Kline, R. (2004). *Principles and practice of structural equation modeling*. New York, United States of America: Guilfor.
- Leyva, D., Weiland, C., Barata, M.C., Yoshikawa, H., Snow, & A., Treviño, E. (2013). Factor structure and predictive validity of the CLASS in prekindergarten: A cross-cultural study across the United States and Chile. Manuscript in preparation.
- Mascareño, M. (2012). Coding Scheme for the Analysis of Moment-by-Moment Teacher-Child(ren). Non published document.
- Mascareño, M., Bosker, R., & Doolaard, S. (2010). Pre-school classrooms: teacher-children interactions that facilitate learning and children's school readiness. Proyecto de tesis para optar al título de Doctor en Educación. University of Groningen, Holanda.
- McCabe, L. A., Rebello-Britto, P., Hernandez, M., & Brooks-Gunn, J. (2004). Games children play: Observing young children's self-regulation across laboratory, home, and school settings. In R. Del Carmen-Wiggins, & A. Carter (Eds.), *Handbook of Infant, Toddler, and Preschool Mental Health Assessment* (pp. 491-521). New York, NY: Oxford University Press.
- McDonald, J. F., & Moffitt, R. A. (1980). The uses of Tobit analysis. *Review of Economics and Statistics*, *62*, 318-321.
- Mendive, S., Weiland, C., & Crosson, A. (2013). Fidelity of implementation in a coaching professional development program in Chile: Relating coaching fidelity to teacher fidelity and to preschoolers' outcomes. Manuscrito en preparación.
- MINEDUC (s/f) Orientaciones para la evaluación de aprendizajes en dominio lector y la comprensión lectora. Recuperado de

- http://www.planesdemejoramiento.cl/m_anexos/Anexo%20IV.pdf el 4 de febrero de 2013.
- MINEDUC (2006). *Conociendo a los dinosaurios. Guía Didáctica 2° Nivel de Transición. Educación Parvularia. Módulo 2*. Santiago: Ediciones MINEDUC. Recuperado de <http://www.mineduc.cl/biblio/documento/200703022116130.NT2MModulo2ProfesorDinosauriosP.pdf> el 4 de febrero de 2013.
- MINEDUC (2008a). *Programa pedagógico. Primer Nivel de Transición*. Santiago: Ediciones MINEDUC.
- MINEDUC (2008b). *Programa pedagógico. Segundo Nivel de Transición*. Santiago: Ediciones MINEDUC.
- Morrison, F. J., Ponitz, C. C., & McClelland, M. M. (2010). Self-regulation and academic achievement in the transition to school. In S. D. Calkins, & M. A. Bell (Eds.), *Child Development at the Intersection of Emotion and Cognition* (1st ed., pp. 203-224). Washington, D.C.: American Psychological Association.
- Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S. & Mather, N. (2005). *Batería III Woodcock-Muñoz*. Itasca, IL: Riverside Publishing.
- Murnane, R., & Willett, J. (2010). *Methods matter: Improving causal inference in educational research*. New York, United States of America: Oxford University Press.
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
- NICHD Early Child Care Research Network (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology*, 41, 428–442.
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, 126(2), 220-246.
- Odom, S., Fleming, K., Diamond, K., Lieber, J., Hanson, M., Butera, G., Horn, E., Palmer, S., Marquis, J. & Children's School Success Project. (2010). Examining different forms of implementation and in early childhood curriculum research. *Early Childhood Research Quarterly*, 25,314-328.
- Phillips, Gormley, & Lowenstein. Classroom Quality and Time Allocation in Tulsa's Early Childhood Programs. Paper presented at the biennial meetings of the Society for Research in Child Development, Boston, MA, March 30, 2007.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *The Classroom Assessment Manual, pre-K*. Baltimore, MD: Brookes.
- Preschool Curriculum Evaluation Research Consortium (2008). *Effects of Preschool Curriculum Programs on School Readiness (NCER 2008-2009)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Washington, DC: U.S. Government Printing Office.
- Preiss, D., Grau, V., Nuñez, M., Alegría, I., Espinoza, & A., Ponce, L. "Instructional patterns remain the same across a segregated educational system: The case of Chile". Paper presented at the 14th Biennial Conference de la European Association for Research on Learning and Instruction, Exeter, UK, Agosto 30, 2011.
- Rimmele, R. (2004). Videograph. Multimedia-Player zur Kodierung von Videos. Kiel: 437 IPN.

- Schochet, P. Z. (2008). *Guidelines for multiple testing in impact evaluations of educational interventions*. Mathematica Policy Research. Unpublished document.
- Shadish, Cook, & Campbell (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.
- Singer, J. & Willett, J. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. New York, United states of America: Oxford University Press.
- Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the preschool self-regulation assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, 22(2), 173-187.
- Snow, C.E., Burns, M.S., & Griffin, P. (1998). (Eds.). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Staub, F. C., West, L. & Bickel, D. D. (2003). What is Content-Focused Coaching? In L. West & F. C. Staub, Content-Focused Coaching. Transforming mathematics lessons (pp. 1-17). Portsmouth, NH: Heinemann.
- Stein, M., Berends, M., Fuchs, D., McMaster, K., Saenz, L., Yen, L., Fuchs, L.S., & Compton, D.L. (2008). Scaling up an early reading program: relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis*, 30, 368-388.
- Storch, S., & Whitehurst, G. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38, 934-947.
- Strasser, K. (2006). Manual para Observación en Sala de Kindergarten. Proyecto Fondecyt N°1040757. Documento no publicado.
- Strasser, K. & Lissi, M. R. (2009). Home and instruction effects on emergent literacy in a sample of Chilean kindergarten children. *Scientific Studies of Reading*, 13(2), 175–204.
- Strasser, K., Lissi, M. R., & Silva, M. (2009). Gestión del tiempo en 12 salas Chilenas de kindergarten: Recreo, colación y algo de instrucción. *Psykhé*, 18, 85-96.
- Treviño, E., Reyes, J., & Godoy Ossa, F. (2010). *Data on UBC comunas and other low-income comunas in Santiago de Chile*. Datos no publicados, Universidad Diego Portales.
- Valencia, E. & Manzi, J. (2011). Desempeño docente: relaciones con antecedentes de los profesores y su contexto. En Manzi, J., Gonzalez, R., & Sun, Y. (Eds.). *La evaluación Docente en Chile*. (pp. 177-193). Santiago de Chile: Centro de Medición MIDE UC.
- Wagner, D.A., Venezky, R.L., & Street, B.V. (1999). *Literacy: An international handbook*. Boulder, CO: Westview Press.
- Wasik, B., & Hindman, A. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology*, 103, 455-469. doi: 10.1037/a0023067
- Weiland, C. (2012). Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills. Cambridge, MA: Doctoral dissertation.

- Werthamer-Larsson, L., Kellam, S., & Wheeler, L. (1991). Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology, 19*(4), 585-602.
- Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. latent structure. *Developmental Psychology, 44*(2), 575-587.
- Woodcock, R. W., Muñoz-Sandoval, A. F., Rief, M., & Alvarado, C. G. (2005). Woodcock Muñoz Language Survey – Revised. Itasca, IL: Riverside Publishing.
- Yoshikawa, H., Leyva, D., Snow, C.E., Treviño, E., Rolla, A., Barata, M.C., Weiland, C., & Arbour, M.C. (2012). *Interim impacts on classroom quality of an initiative to improve the quality of preschool education in Chile: A cluster-randomized trial*. Manuscrito en preparación.
- Yoshikawa, H., Barata, M.C., Rolla, A., Da Silva, C.D., Ayoub, C., Arbour, M. & Snow, C.E. (2008, forthcoming). Un Buen Comienzo, una iniciativa para mejorar la educación preescolar en Chile: Resultados del primer año de implementación. Report to the UNICEF office, Santiago, Chile.
- Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nat. Protocols, 1*(1), 297-301.

Anexo 1*

Selección de municipalidades en Santiago de Chile. Para ser invitado a participar en el estudio UBC, cada municipalidad tuvo que cumplir los siguientes criterios: a) tener una alta proporción de niños en condición de vulnerabilidad; b) tener un mínimo de ocho escuelas con aulas de prekinder y kínder; y c) estar ubicadas en la Región Metropolitana de Santiago. Los niños en condición de vulnerabilidad fueron identificados utilizando los criterios del Ministerio de Educación de Chile, que incluyó ingreso familiar, educación de los padres, y si la familia recibe beneficio social o de salud del gobierno. Otras características de la municipalidad tales como proximidad y accesibilidad, seguridad, ausencia de otras intervenciones llevadas a cabo en el nivel preescolar en el mismo período (particularmente de desarrollo de lenguaje), administración local de clínicas, y presencia de Programa IRA (Infecciones Respiratorias Agudas) en los centros de salud presentes en la comuna fueron considerados cuando se seleccionaron las municipalidades.

Catorce municipalidades cumplieron estos criterios y fueron invitadas a postular para participar en el programa UBC (una municipalidad en cohorte 1, seis municipalidades en cohorte 2 y siete municipalidades en cohorte 3). Se condujeron entrevistas con autoridades municipales, incluyendo representantes de los alcaldes y de departamentos de educación y salud, para clarificar los objetivos, explicar el diseño de evaluación, y responder preguntas. En total, seis municipalidades fueron seleccionadas (una en cohorte 1, dos en cohorte 2, y tres en cohorte 3). Una lotería pública se llevó a cabo para asignar aleatoriamente las escuelas al interior de cada municipalidad al grupo de condición intervención UBC (Intervención I) o la condición de comparación (Intervención II; ver más abajo). En total, 32 escuelas (53 aulas) fueron asignadas a la condición de Intervención I y 32 escuelas (39 aulas) fueron asignadas a la condición de Intervención II. Las aulas de la misma escuela fueron siempre asignadas a la misma condición, para evitar contaminación de la condición al interior de las escuelas (ej., adopción de aspectos de UBC en la condición de Intervención II; Murnane & Willett, 2010). Los 25 consultorios en las municipalidades seleccionadas participaron en UBC. Cada uno entrega servicios de salud universal, según lo exigido por la ley chilena, a los niños en las escuelas de los grupos de Intervención 1 e Intervención 2. Sin embargo, el foco de la intervención de salud de UBC, que pretendió proveer vínculos más estrechos entre las escuelas, familias y los consultorios, fueron las escuelas. Contaminación entre condiciones no fue una amenaza mayor al componente de salud porque la escuela, no el consultorio fue el foco de la intervención.

Las seis municipalidades que participaron en el estudio no difirieron de aquellas elegibles pero que no fueron seleccionadas en puntajes SIMCE o en el índice de vulnerabilidad socioeconómica (Treviño, Reyes, y Godoy, 2010).

Reclutamiento de las escuelas, educadoras y familias. Todos los representantes de los departamentos de educación de las municipalidades involucraron a directores de las escuelas que satisficieron los criterios del estudio. Las educadoras participaron en el estudio por requerimiento de los directivos de sus escuelas.

El consentimiento de los padres para participar en el estudio y la información sociodemográfica de los cuestionarios, de la conducta infantil, de prácticas de salud y de alfabetización inicial, fueron recolectadas en las escuelas, en encuentros de grupos pequeños, o en su defecto en el hogar, por el equipo de evaluadores. Los evaluadores entrenados administraron los cuestionarios y leyeron cuidadosamente las preguntas cuando los padres que tenían dificultades para leer así lo requerían. Si los padres no sabían escribir, los evaluadores escribieron sus respuestas por ellos.