



Programa Crecer Jugando
Evaluación de Impacto - Reporte Final

Mayo 2025



Este proyecto de investigación fue financiado por la Education Endowment Foundation, en colaboración con la BHP Foundation, en el marco del proyecto "Building a global evidence ecosystem for teaching"(Acuerdo de subvención EEF-032022-SUMMA), a través del Fondo Impacto.

La evaluación contó además con el apoyo financiero de Fundación Angelini, Fundación Colunga, Fundación Hue-lla, Fundación Viento Sur, Fundación Ilumina y Fundación Olivo.

Las opiniones expresadas en este trabajo son de responsabilidad de los autores y no representan necesariamente las del Education Endowment Foundation – EEF, BHP Foundation o de sus respectivas Juntas Directivas.

Equipos de Trabajo:

SUMMA: Laboratorio de Investigación e innovación en educación para América Latina y el Caribe

- María José Sepúlveda, Directora de Evaluación
- Paula Salas, Coordinadora de Evaluación
- Nicolás Caglevic, Analista de Evaluación

Equipo Estudios y Consultorías Focus

- Isidora Fuenzalida, Jefa de Proyecto
- Benjamín Coloma, Coordinador Metodológico
- Juan Pablo Madrid, Coordinador de Terreno
- Benjamin Nielsen, Coordinador de análisis PICCOLO

Equipo Infancia Primero

- María de los Ángeles Castro, Directora Ejecutiva
- Magdalena Mongillo, Directora Social
- Daniela Nuñez, Investigadora
- María José Osimani, Coordinadora de Operaciones

Investigadores: María Jesús Medina, Verónica Acevedo e Ignacia Arredondo.

Analistas PICCOLO: Catalina Undurraga, Diana Gomez y Camila Pinto.

Cómo citar el documento: SUMMA, Estudios y Consultorías Focus (2025). *Programa Crecer Jugando. Evaluación de Impacto*. Proyecto Fondo Impacto "Construyendo un ecosistema global de evidencia para la mejora de la enseñanza".

Resumen ejecutivo

El proyecto

- *Descripción y objetivo:* Crecer Jugando es un programa preventivo y promocional dirigido a fortalecer el vínculo y la calidad de la interacción entre cuidadores/as y niños/as de 0 a 48 meses. A través del desarrollo de habilidades parentales y la promoción de interacciones cotidianas (como jugar, leer o cantar), busca favorecer el desarrollo infantil.
- *Estrategia de intervención:* El programa contempla 10 sesiones grupales semanales con contenidos prácticos, complementadas por un foro digital de apoyo. Se implementa en modalidad presencial y *online*, manteniendo ambas opciones tras la pandemia por COVID-19.
- *Población beneficiaria:* Díadas cuidador/a–niño/a vinculadas a servicios públicos y que viven en contextos de alta vulnerabilidad social.
- *Implementación:* Ejecutado por duplas de facilitador/a profesional y monitor/a, bajo coordinación de instituciones locales. Fundación Infancia Primero entrega capacitación, supervisión y acompañamiento técnico.

Evaluación de impacto

- *Diseño metodológico:* Evaluación de efectividad bajo condiciones reales, con diseño experimental y asignación aleatoria de 600 díadas. Se estimaron impactos sobre resultados primarios (vínculo y desarrollo infantil), intermedios (habilidades parentales y frecuencia de interacción) y secundarios (estrés parental y apoyo social). Finalmente participaron 324 díadas en la evaluación.
- *Métodos de estimación:* El análisis incluyó estimaciones ITT y LATE, efectos marginales y costo-efectividad. Se realizaron pruebas de balance y robustez para asegurar la validez de los resultados y controlar sesgos potenciales.
- *Levantamiento de información y aplicación de instrumentos:* El trabajo de campo se realizó entre marzo de 2024 y enero de 2025. La aplicación de instrumentos se realizó tanto al grupo de tratamiento como al grupo de control, al término del ciclo de intervención correspondiente a cada grupo.

Aportes generales de la evaluación

- El estudio aporta evidencia de la efectividad del programa en línea con su modelo de cambio, considerando resultados intermedios, primarios y secundarios.
- La evaluación entrega información sobre su efectividad considerando edad del infante, modalidad e intensidad de participación, e identifica umbrales mínimos de asistencia.
- Se incorporan análisis de costo-efectividad relevantes para decisiones sobre escalabilidad.

Resultados de impacto

- Mejoras significativas en habilidades parentales (E2P), especialmente en dimensiones reflexiva y formativa.
- Aumento significativo en la frecuencia de interacciones (leer cantar, jugar) entre las díadas que asistieron al menos a 4 sesiones.
- Reducción del estrés parental (PSI-4-SF) sobre 0.4 DE en todas las dimensiones.

- Impacto positivo en desarrollo infantil (ASQ-3) en el dominio socio individual; sin efectos en otras áreas, posiblemente por la proximidad temporal entre intervención y medición.
- Mejora en vínculo cuidador–niño/a en la dimensión enseñanza (PICCOLO), aunque con posible sesgo por autoselección.

Interpretación de efectos (ITT y LATE)

- La estimación ITT muestra efectos positivos, reflejando el impacto general del programa en condiciones reales de implementación. El análisis LATE, centrado en quienes participaron efectivamente, arroja efectos más altos en todos los indicadores.
- Ambos estimadores ofrecen una comprensión complementaria: mientras el ITT refleja la efectividad promedio del programa en la población asignada al tratamiento, el LATE estima la efectividad bajo condiciones óptimas de participación.

Efectos según intensidad de participación

- En promedio, los efectos del programa se consolidan y comienzan a ser significativos a partir de la asistencia a la tercera sesión.
- Esta información permite diseñar estrategias focalizadas de retención para diadas cercanas al umbral mínimo para alcanzar impacto.

Efectos diferenciados

- El Grupo Exploradores (13–24 meses) muestra mayores mejoras en habilidades parentales, especialmente en las dimensiones reflexiva y protectora, así como en el desarrollo de la comunicación, lo que sugiere una mayor sensibilidad del programa a las necesidades de esta etapa del desarrollo.
- En relación al vínculo cuidador/a-infante, se observa que el programa impacta en mayor medida al grupo acurrucados tanto en las dimensiones enseñanza como responsividad.
- La reducción del estrés parental es mayor en los grupos Exploradores y Creadores, posiblemente por las mayores exigencias y desafíos propios de esas etapas, en comparación con el grupo Acurrucados (0–12 meses).
- En cuanto a las modalidades de implementación, no se observaron diferencias significativas entre presencial y *online*, salvo en apoyo social percibido (mayor en presencial). La menor potencia estadística del grupo *online* limita conclusiones más específicas.

Costo-efectividad

- El costo promedio de implementación por diada es de aproximadamente \$137.000 CLP.
- Los ratios de costo-efectividad se sitúan entre \$498.000 y \$600.000 por desviación estándar, cifras que resultan competitivas frente a programas internacionales comparables.
- En conjunto, el programa presenta un modelo eficiente, lo que fortalece su sostenibilidad económica y potencial de escalabilidad.

Proyección y desafíos

- Se recomienda hacer seguimiento longitudinal para observar de mayor plazo, especialmente en desarrollo infantil, y ajustar estrategias en función de los aprendizajes observados.

Contenido

1. Introducción	7
2. Descripción del programa y fundamentos de la evaluación	8
2.1. Contexto y relevancia de la evaluación	8
2.2. Descripción del programa Crecer Jugando y su teoría de cambio	9
2.3. Resultados Esperados	10
3. Objetivos de la evaluación de impacto	11
3.1. Objetivo general:	11
3.2. Objetivos específicos:	11
4. Diseño Metodológico	12
4.1. Diseño de evaluación	12
4.2. Universo y criterios de inclusión de participantes	12
5. Instrumentos de medición	14
5.1. Instrumentos por tipo de resultado	14
5.2. Variables de control	15
6. Métodos de estimación	16
6.1. Intención de Tratamiento (ITT)	16
6.2. Límite Local de Tratamiento Promedio (LATE)	17
6.3. Efectos heterogéneos	17
6.4. Efectos marginales	18
6.5. Costo-efectividad	19
6.6. Robustez	19
7. Trabajo de campo	20
7.1. Procedimiento de aleatorización y contacto	20
7.2. Nivel de respuesta y ajustes por atrición	21
8. Cálculos Efecto Mínimo Detectable	27
9. Resultados	29
9.1. Caracterización de los participantes	29
9.2. Fidelidad de la implementación	31
9.2.1. Participación por asistencia a sesiones	31
9.2.2. Participación vía WhatsApp	32
9.2.3. Participación efectiva en el programa	32
9.3. Resultados de impacto	32
9.3.1. Resultados intermedios	33
9.3.2. Resultados primarios	34
9.3.3. Resultados secundarios	34
9.4. Efectos marginales	37
9.5. Efectos heterogéneos	46
9.5.1. Efectos heterogéneos sobre modalidad de tratamiento	46
9.5.2. Efectos heterogéneos sobre grupo etario asignado	52
9.6. Análisis de robustez: sesgo por atrición	58

10. Análisis de costo efectividad	62
10.1. Costos de la intervención	62
10.2. Ratio de costo efectividad	63
10.3. Estrés Parental (PSI-4)	67
10.4. Habilidades parentales (E2P)	67
11. Conclusiones y desafíos	70
12. Referencias Bibliográficas	74

1 Introducción

Este informe presenta los resultados finales de la evaluación de impacto del programa Crecer Jugando, desarrollada entre 2023 y 2024 por encargo de Fundación Infancia Primero. La evaluación se enmarca en el Fondo Impacto EDU-LAC, una iniciativa impulsada por SUMMA en colaboración con la Education Endowment Foundation (EEF) y financiada por la Fundación BHP, como parte de su Programa Global de Equidad Educativa. Esta iniciativa busca generar nueva evidencia sobre intervenciones educativas innovadoras y contribuir a la mejora de políticas y prácticas en América Latina y el Caribe. En este contexto, el programa Crecer Jugando fue seleccionado como una de las intervenciones a evaluar, siendo Fundación Infancia Primero responsable de su implementación y Focus del diseño y ejecución de la evaluación de impacto y procesos, bajo estándares de calidad y principios de colaboración, independencia y uso de la evidencia como bien público.

La evaluación tuvo como propósito estimar los efectos del programa en sus resultados esperados, definidos en su modelo de cambio: el fortalecimiento de competencias parentales, la mejora en el vínculo entre cuidadores y niños/as, y la promoción del desarrollo infantil en sus primeras etapas. Asimismo, se incorporó el análisis del estrés parental como resultado secundario, dada su influencia sobre la calidad de la interacción cuidador-niño/a.

El estudio fue diseñado como una evaluación de impacto con asignación aleatoria, lo que permitió estimar efectos causales a partir de un diseño experimental. La intervención fue implementada en formato presencial y remoto (*online*) en distintas comunas del país, lo que permitió explorar su efectividad bajo diferentes condiciones de ejecución. Además de los análisis principales, se incorporaron estimaciones complementarias, como un estimador Local Average Treatment Effect (LATE) y análisis de efectos marginales.

La evaluación también incluyó análisis de efectos heterogéneos y una estimación de costo-efectividad, con el objetivo de generar evidencia útil para la toma de decisiones sobre ajustes al diseño, condiciones de escalabilidad y sostenibilidad del programa. En su conjunto, los resultados entregan una mirada integral sobre el desempeño del programa, sus alcances inmediatos y los desafíos que abre hacia futuras etapas de seguimiento y consolidación.

2 Descripción del programa y fundamentos de la evaluación

2.1 Contexto y relevancia de la evaluación

La evidencia científica subraya la importancia de la primera infancia para el desarrollo cerebral y la adquisición de habilidades fundamentales para el futuro. Diversas disciplinas —como la educación, la neurobiología y la economía— coinciden en que el desarrollo humano en los primeros años de vida sienta las bases para el desempeño en etapas posteriores (Shonkoff, 2000; Heckman, 2000). En particular, se ha documentado que las experiencias tempranas inciden en la estructura y funcionalidad del cerebro (Fox, Levitt y Nelson, 2010). El ambiente en el que crecen los niños y niñas es un factor central para su desarrollo, existiendo una desventaja evidente para aquellos que viven en contextos de mayor vulnerabilidad. Asimismo, se ha documentado que la crianza, entendida como el entorno de cuidado inmediato, constituye el primer proceso de socialización y un componente determinante en el desarrollo infantil temprano. La calidad de estas experiencias iniciales es central para que los niños y niñas puedan alcanzar su máximo potencial de desarrollo (Fox, Levitt y Nelson, 2010).

Crecer Jugando es un programa de intervención temprana para familias con niños o niñas menores de 4 años que busca instalar un modelo de apoyo a la parentalidad y desarrollo infantil temprano, con el fin de fortalecer el vínculo y la interacción entre el niño o niña y su cuidador/a.

El programa Crecer Jugando está respaldado teóricamente a partir de evidencia internacional, además de contar con evidencia propia en diversas fases de su instalación. En su actual fase de desarrollo, se busca llevar adelante un proceso de alta rigurosidad, que sirva como insumo para lograr un impacto positivo a gran escala, buscando las mejores estrategias que den cuenta de su efectividad para resolver los problemas que dan origen a su existencia y que, a pesar de los años, siguen igualmente vigentes.

A la fecha, el programa ha obtenido buenos resultados en cuanto a su implementación y adherencia, considerando las comunas donde se ejecuta actualmente, lo que ha permitido fortalecer los componentes de su estructura. Las evaluaciones realizadas hasta a la fecha demuestran que el programa logra cumplir sus objetivos, contando con impactos estadísticamente significativos en varios de sus indicadores de medición.

Desde el diseño del programa el año 2014 y la ejecución de su piloto el año 2015, el equipo de Infancia Primero ha llevado a cabo diversos procesos de evaluación, utilizando distintas metodologías, para verificar el diseño, la implementación y los resultados de la intervención de manera periódica. Todos estos procesos han implicado la incorporación de mejoras y actualizaciones que permiten al programa mantenerse vigente y responder de manera adecuada a las diferentes necesidades de las familias con quienes trabaja.

Específicamente, se han realizado 5 evaluaciones del Programa Crecer Jugando con equipos externos y expertos/as en la materia, que han permitido analizar los diversos procesos de cambio que ha experimentado. Estos resultados han dado cuenta de una implementación fiel a su diseño, con una adecuada capacitación y aplicación del modelo de transferencia a los equipos ejecutores, y resultados positivos (pre-post) utilizando escalas de evaluación propias que mostraron confiabilidad y validez (fortalecimiento de redes de apoyo, aumento de conocimientos y habilidades de buen trato, frecuencia de interacción entre cuidadores y niños/as). Toda esta información se encuentra disponible en la página web de la Fundación (infanciaprimer.org).

Con la evidencia acumulada sobre la correcta implementación en procesos de formación y transferencia del modelo, así como resultados pre y post en participantes que se mantienen en el programa, surge la necesidad de avanzar en evaluar el impacto, de manera de complementar y profundizar el conocimiento generado en las evaluaciones anteriores para validar los resultados alcanzados y conocer si estos son atribuibles al programa. Además, la evaluación de impacto proporcionará información para respaldar decisiones estratégicas relacionadas

con su escalamiento y sostenibilidad, junto con proporcionar evidencia para el fortalecimiento de otros programas similares por el conocimiento que generan los análisis complementarios como el de costo-efectividad.

Con este objetivo, se solicita a Estudios y Consultorías Focus realizar un estudio que permita identificar si la implementación del Programa Crecer Jugando genera algún impacto en sus participantes en los resultados intermedios, finales y secundarios definidos en su teoría de cambio, en el marco del convenio tripartito IMPACTO EDU-LAC para el desarrollo de Evaluaciones de Impacto en América Latina y el Caribe de SUMMA.

2.2 Descripción del programa Crecer Jugando y su teoría de cambio

Crecer Jugando es un programa orientado al apoyo en la parentalidad y al desarrollo infantil temprano, cuyo objetivo es fortalecer el vínculo y la calidad de la interacción entre niños y niñas y sus cuidadores principales (díadas), como mecanismo para favorecer el desarrollo infantil. Para ello, el programa entrega competencias parentales clave y busca fortalecer las redes de apoyo entre cuidadores. Su población objetivo está compuesta por niños y niñas de entre 0 y 48 meses y sus cuidadores..

Es un programa preventivo y promocional, basado en evidencia, que combina contenido teórico y práctico para los cuidadores buscando fomentar la sensibilidad, la capacidad de reflexión, conocimientos y sus habilidades para contener y regular emocionalmente a sus hijos/as. En términos generales, la intervención se enfoca en proporcionar información a los cuidadores de niños menores de 48 meses sobre aspectos clave de la crianza, fortaleciendo sus habilidades parentales y estimulando la reflexión acerca de sus prácticas y desafíos en la crianza.

El programa está organizado en ciclos que consisten en 10 sesiones de talleres grupales en donde participan un grupo de díadas. Los talleres se dan de forma semanal y cada ciclo tiene una duración aproximada de 3 meses. Cada sesión tiene una duración de 1 hora y media. Las sesiones se realizan en una sala equipada, que cumple con ciertas condiciones establecidas por el programa. Esto se basa en el principio de un ambiente preparado, que fomente el juego, el desarrollo, el vínculo y la interacción.

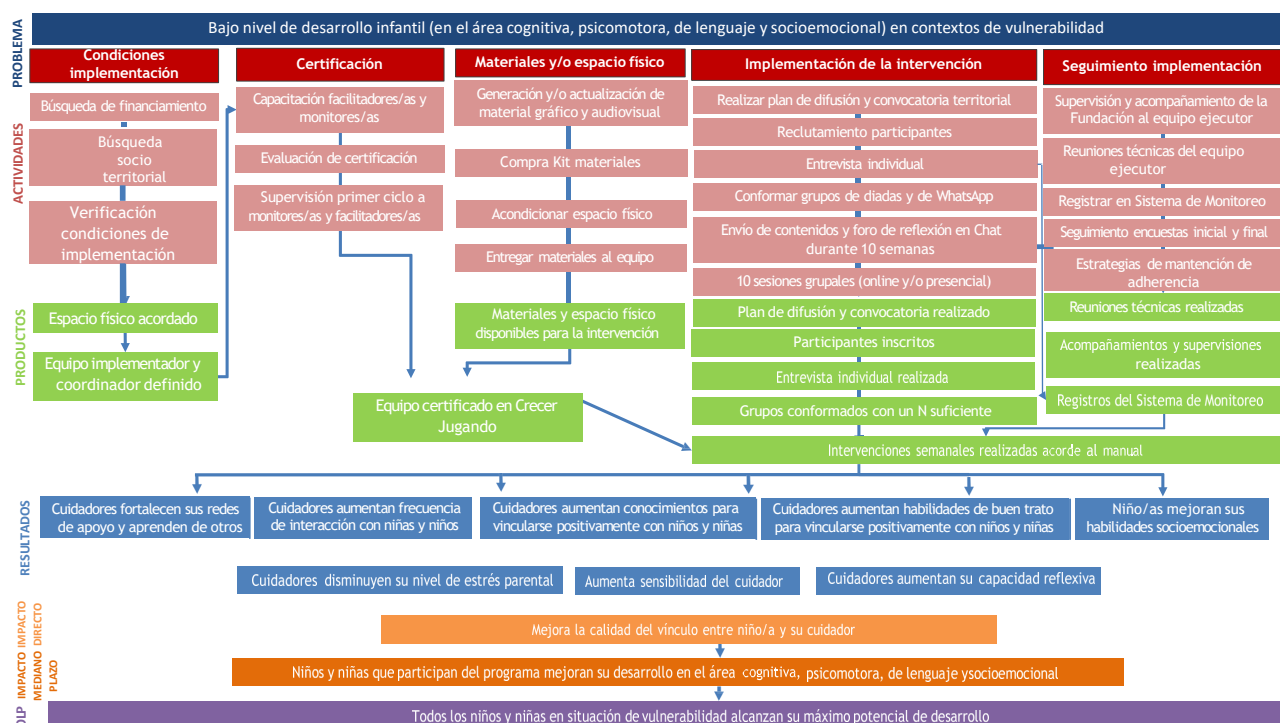
Además de las sesiones, los cuidadores participan de un foro comunitario a través de plataformas tecnológicas a través de la cual reciben información adicional referente al contenido desarrollado en cada sesión, cuentan con la oportunidad de conectar entre sí y resuelven dudas relacionadas al programa y su contenido.

Dada la naturaleza de prevención y promoción del programa, Crecer Jugando se instala en instituciones locales del Estado, ya sea en los sectores de salud, educación u otros. Esto se debe a que la población objetivo acude a estos espacios para obtener servicios que acompañan la trayectoria de niños y niñas desde su gestación hasta el inicio del ciclo escolar.

La pandemia de 2019 impulsó la necesidad de adaptar el programa a un formato virtual, permitiendo su continuidad durante las fases de confinamiento. Aunque muchas comunas regresaron a la modalidad presencial una vez levantadas las restricciones, en algunas áreas se vio en el formato *online* una oportunidad para ampliar el alcance del programa y llegar a familias en zonas más remotas. Por ello, el modelo siguió adaptándose, y hoy el programa Crecer Jugando se ofrece tanto en modalidad presencial como *online*.

En cuanto a los actores que forman parte de la implementación, el programa involucra a dos actores claves: la contraparte local y el equipo ejecutor. La contraparte local tiene la responsabilidad de designar y coordinar al equipo ejecutor, mientras que este último se encarga de realizar las intervenciones directamente con las familias. En relación a la conformación del equipo ejecutor, este debe estar integrado por un facilitador profesional y un monitor remunerado, ambos con una disponibilidad mínima de 16 horas semanales. Por último, la capacitación es llevada a cabo por la Fundación Infancia Primero, cuyo rol también es supervisar y acompañar de manera constante las actividades desarrolladas por ambos actores.

Figura 1: Teoría de Cambio



La Teoría de Cambio del programa describe de manera causal los resultados esperados (resultados primarios) a nivel intermedio y final. El programa está diseñado para que, a través de sus actividades, los cuidadores se vean expuestos a la intervención y, por lo tanto, logren obtener resultados por su participación. (Figura 1)

2.3 Resultados Esperados

El programa busca alcanzar los siguientes resultados intermedios en los/as cuidadores/as:

- **Fortalecimiento de habilidades parentales:** Los cuidadores mejoran sus habilidades para criar y educar a sus hijos.
- **Incremento de la interacción con los niños y niñas:** Los cuidadores incrementan la frecuencia de actividades como jugar, cantar y leer con sus hijos e hijas, haciéndolo más frecuentemente que antes de participar en el programa.

La medición de estos resultados es relevante, dado que la literatura ha demostrado que actividades simples como cantar, jugar o leer un cuento pueden tener un impacto significativo en el desarrollo de los niños, mejorando sus habilidades para pensar, comunicar y conectar con otros. A pesar de esto, muchos padres, madres y cuidadores no son conscientes de la importancia de estas acciones y, por lo tanto, no las realizan (Nores y Barnett, 2009; Maulik y Darmstadt, 2009).

Luego de generarse los cambios en los resultados intermedios, el modelo de cambio del programa identifica dos resultados primarios: mejorar el **vínculo entre el/la cuidador/a y el niño o niña**, lo que a su vez, debería llevar a mejoras en el **desarrollo de los infantes** en áreas como la psicomotricidad, el lenguaje, las habilidades cognitivas y socioemocionales, identificadas como los resultados primarios finales del programa.

Múltiples investigaciones han establecido que uno de los predictores más importantes y consistentes del desarrollo social y cognitivo es la calidad de la interacción entre el cuidador y el niño. Las interacciones más sensibles, responsivas, atentas y cognitivamente estimulantes por parte del cuidador están asociadas con mejores resultados en el desarrollo del niño. Estos hallazgos incluyen la seguridad del apego, el desarrollo del lenguaje, las habilidades numéricas y de letras, y el comportamiento social (NICHD, 2006).

Finalmente, el diseño del programa permite identificar resultados secundarios que están relacionados con los resultados esperados. Estos resultados han sido identificados como secundarios ya que, si bien apoyan el logro del objetivo central, no son los indicadores primarios del éxito del programa.

- **Fortalecimiento de redes de apoyo:** Los cuidadores fortalecen sus redes de apoyo en la crianza y aprenden de otros. Su relación con el objetivo principal es indirecta, ya que estas redes contribuyen a mejorar la calidad de la crianza al proporcionar apoyo emocional y compartir experiencias, pero no son el foco central del programa.
- **Reducción del estrés parental:** Los cuidadores experimentan una disminución en su nivel de estrés parental. La reducción del estrés parental es un efecto colateral positivo del programa, ya que un menor estrés contribuye a mejorar la interacción entre la madre y el hijo.

3 Objetivos de la evaluación de impacto

3.1 Objetivo general:

Evaluar el impacto del programa Crecer Jugando en el fortalecimiento del vínculo entre los cuidadores principales y los niños y niñas de 0 a 48 meses, así como en el desarrollo infantil temprano, a través de la mejora de las habilidades parentales, el aumento de la frecuencia de sus interacciones, la disminución del estrés parental y el fortalecimientos de las redes de apoyo.

3.2 Objetivos específicos:

1. Evaluar el impacto del programa Crecer Jugando en sus resultados primarios: fortalecimiento del vínculo cuidador/a-niño/a y desarrollo infantil
2. Evaluar el impacto del programa Crecer Jugando en sus resultados intermedios: habilidades parentales y frecuencia de interacciones entre cuidadores/as y niños/a
3. Evaluar el impacto del programa Crecer Jugando en sus resultados secundarios: reducción del estrés parental y fortalecimiento de redes de apoyo

4 Diseño Metodológico

4.1 Diseño de evaluación

El programa Crecer Jugando fue evaluado mediante un diseño experimental, que contempló la asignación aleatoria de las diadas participantes (compuestas por un/a niño/a y su cuidador/a principal) a dos grupos: uno de tratamiento, que accedió a la intervención, y uno de control, que no la recibió. Esto busca asegurar que la única diferencia entre los grupos sea la exposición al programa, lo cual permite atribuir los cambios observados al impacto de la intervención, fortaleciendo la validez interna de la investigación.

4.2 Universo y criterios de inclusión de participantes

El programa se implementa en 14 comunas del país, pero la investigación se realizó en las siguientes 8 comunas: El Bosque, Melipilla, Quinta Normal, Limache, Concepción, Maule, Monte Patria y Río Negro. Estas comunas cuentan con equipos estables y capacitados en los territorios, de acuerdo con el modelo del programa.

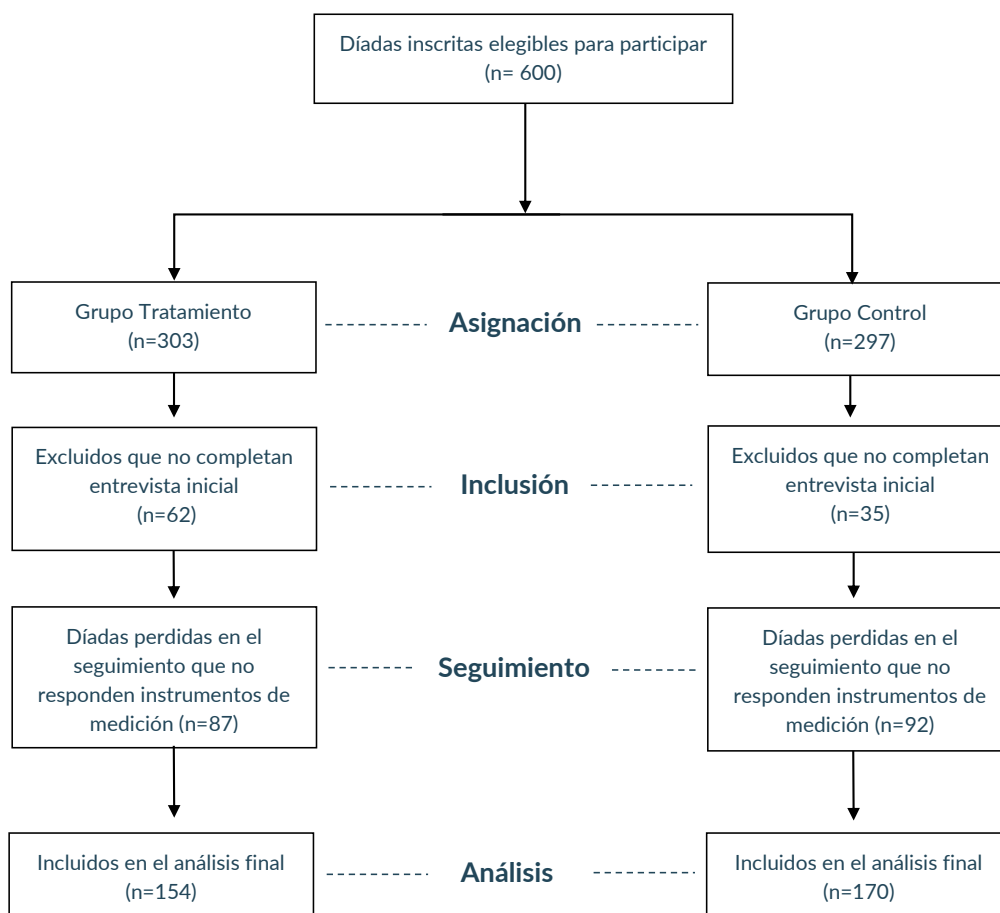
Las diadas seleccionadas para la aleatorización debían cumplir con los siguientes criterios:

1. Inscribirse voluntariamente en el programa.
2. No haber participado en ediciones anteriores del programa
3. No recibir ninguna otra intervención relacionada con el desarrollo de habilidades parentales.
4. No tener un diagnóstico asociado al desarrollo psicomotor que pudiera afectar su participación en el programa.
5. Inscribirse antes de que se completara el número máximo de diadas atendibles en el ciclo correspondiente. Las diadas inscritas después de alcanzar este límite no fueron elegibles para ese ciclo, pero podían considerarse en futuros ciclos si mantenían su interés.

A modo de resumen, en la Figura 2 se observa el proceso mediante el cual se logra la muestra final de análisis.

Cabe destacar que la muestra utilizada en el estudio consistió en diadas que participaron voluntariamente en el programa, recurriendo a un muestreo no aleatorio. Debido a esta metodología de selección, existe un sesgo inherente en las características de las diadas que conforman el universo de estudio. Este sesgo limita la validez externa de los resultados, lo que significa que los efectos identificados no pueden ser generalizados a otros grupos de diadas fuera de esta muestra específica.

Figura 2: Diagrama flujo de evaluación



Fuente: elaboración propia

5 Instrumentos de medición

5.1 Instrumentos por tipo de resultado

A continuación, se describen los instrumentos de medición utilizados para evaluar los resultados intermedios, primarios y secundarios del estudio.

Resultados Intermedios

- **Habilidades parentales:** se utilizó la Escala de Parentalidad E2P, desarrollada en 2011 por Esteban Gómez y Magdalena Muñoz y ajustada a su última versión en 2019, que mide prácticas parentales efectivas de cuidado para el desarrollo de un niño y adolescente, desde los 0 a 17 años de edad, organizada en rangos etarios (Gómez & Contreras, 2019). Dichas prácticas se analizan dentro de un modelo que distingue cuatro competencias parentales: vinculares, formativas, protectoras y reflexivas, cada una conformada a su vez por cinco componentes. El instrumento consta de 60 ítems aplicables a distintos rangos de edad, con cinco opciones de frecuencia, donde “Nunca” asigna 0 puntos y “Siempre” 4 puntos. Para niños menores de 3 meses, las opciones “Nunca”, “Casi Nunca” o “A veces” asignaron 2 puntos, “Casi Siempre” 3 puntos, y “Siempre” 4 puntos, tal como indica su manual. Este instrumento fue aplicado tanto al grupo tratamiento como control, post intervención.
- **Frecuencia de interacción con los niños y niñas:** se utilizó la encuesta inicial y final del programa Crecer Jugando, la que es respondida de manera autoaplicada al inicio y al final del ciclo de intervención. Este instrumento cuenta con 16 ítems para evaluar 4 dimensiones, las que han sido validados a través de análisis de consistencia y confiabilidad por parte de la Fundación (infanciaprimer.org). La escala para frecuencia de interacción (cantan, juegan y leen) consta de 6 ítems de escala Likert con valores numéricos entre 0 y 4. Donde 0 es “Ninguna vez de la última semana” y 4 es “Todos los días de la última semana”. Este instrumento fue aplicado tanto grupo tratamiento como control, pre y post intervención.

Resultados Primarios

- **Vínculo entre el/la cuidador/a y el niño o niña:** se utilizó el instrumento *Parenting Interactions with Children: Checklist of Observations Linked to Outcomes* (Piccolo), lista de ítems de observación relacionado a la interacción cuidador-niño/a - que cuenta con un listado de 29 comportamientos observables en la parentalidad con niños/as de edades entre los 10 y 47 meses orientada al desarrollo. Se estructura en cuatro dimensiones: (1) afecto, (2) responsividad, (3) aliento o ánimo, y (4) enseñanza. Cada dimensión se evalúa mediante ítems calificados de 0 a 2, con un puntaje máximo de 14 para las tres primeras dimensiones y de 16 para la dimensión de enseñanza (Roggman, Cook, Innocenti, Jump Norman, & Christiansen, 2013). Este instrumento se aplica mediante la grabación de una sesión de juego de al menos 10 minutos entre el/la cuidador/a y el infante. El análisis es realizado por personal calificado y certificado, con un tiempo promedio de análisis de aproximadamente 2 horas. Este instrumento fue aplicado tanto grupo tratamiento como control, post intervención.
- **Desarrollo del niño o niña:** se utilizó el (*Ages and Stages Questionnaire* (ASQ-3) que a partir de una escala verifica que el neurodesarrollo de los niños sea normal para su edad. Esta herramienta evalúa las principales áreas de desarrollo, y entrega resultados en cinco dominios: (1) comunicación, (2) motricidad gruesa, (3) motricidad fina, (4) resolución de problemas, y (5) habilidades personales y sociales. Cada ítem se responde con “sí” (puntuación = 10), “a veces” (puntuación = 5), o “todavía no” (puntuación = 0). Cada dominio contiene 6 ítems, por lo que el puntaje total varía entre 0 y 60 por dominio (Squires et al., 2009). Este instrumento fue aplicado tanto grupo tratamiento como control, post intervención.

Resultados Secundarios

- **Redes de apoyo:** se utilizó la encuesta inicial y final del programa Crecer Jugando, a través de la escala de apoyo social percibido, que se construye con 4 ítems cuyas categorías de respuestas toman valores entre 0 y 4. Donde 0 es “Muy en desacuerdo” y 4 “Muy de acuerdo”
- **Estrés parental:** se utilizó el *Parental Stress Index versión 4 Short Form (PSI4-SF)* - Índice de estrés parental -: es una versión abreviada de 36 ítems del Índice de estrés parental, cuarta edición (PSI-4) completo y se divide en tres dimensiones: (1) Angustia parental, (2) Interacción Disfuncional Padre-Hijo, y (3) Características Difíciles del Niño. Cada dimensión consta de 12 ítems, con un puntaje máximo de 60.

Los instrumentos utilizados en esta evaluación han sido previamente adaptados y validados para su uso en población chilena, lo que respalda su pertinencia cultural y psicométrica en el contexto local. En el caso del ASQ-3, su uso ha sido ampliamente difundido en programas públicos como Chile Crece Contigo, y se encuentra validado en Chile desde 2013 en el marco de evaluaciones de desarrollo infantil temprano (Squires et al., 2009). La Escala de Parentalidad Positiva (E2P) fue específicamente desarrollada y validada en población chilena, con un proceso riguroso de revisión teórica y empírica, incluyendo análisis psicométricos en distintas regiones del país (Gómez & Contreras, 2019). Por su parte, la escala PICCOLO ha sido aplicada en diversos estudios en Chile y su uso ha sido promovido por instituciones académicas como el Centro CUIDA de la Pontificia Universidad Católica de Chile, respaldando su validez en el contexto local (Centro CUIDA, 2022). Finalmente, el Parenting Stress Index Short Form (PSI4-SF) cuenta con estudios de validación en población chilena que evidencian adecuados niveles de confiabilidad y validez para medir estrés parental en contextos similares (Aracena et al., 2013).

5.2 Variables de control

Todas las variables de caracterización provienen de la entrevista inicial aplicada a la totalidad de las díadas aleatorizadas. Adicionalmente, se incorpora como variable de control el puntaje obtenido en la encuesta inicial. Estas variables se incluyen en los modelos con el objetivo de controlar por diferencias iniciales entre los grupos tratamiento y control, en caso de que existan desequilibrios estadísticamente significativos en los indicadores medidos al inicio de la intervención.

- **Frecuencia de interacción con los niños y niñas:** se utilizó la encuesta inicial del programa Crecer Jugando, a través de la escala de frecuencia de interacción, que se construye con 6 ítems cuyas categorías de respuestas toman valores entre 0 y 4, donde 0 es “Ninguna vez de la última semana” y 4 es “Todos los días de la última semana”. Cabe destacar que esta variable también es utilizada como variable de resultado intermedio a través de una metodología de diferencias en diferencias.
- **Redes de apoyo:** se utilizó la encuesta inicial del programa Crecer Jugando, a través de la escala de apoyo social percibido, que se construye con 4 ítems cuyas categorías de respuestas toman valores entre 0 y 4, donde 0 es “Muy en desacuerdo” y 4 “Muy de acuerdo”. Al igual que el caso anterior esta variable es utilizada tanto como variable de resultado como de control.
- **Conocimientos y habilidades de buen trato:** se obtiene mediante la encuesta inicial del programa Crecer Jugando, y se construye mediante 5 ítems cuyas categorías de respuestas toman valores entre 0 y 4, donde 0 es “Muy en desacuerdo” y 4 “Muy de acuerdo”.
- **Dificultades parentales:** se obtiene mediante la encuesta inicial del programa Crecer Jugando, y se construye mediante 4 ítems cuyas categorías de respuestas toman valores entre 0 y 4, donde 0 es “Muy en desacuerdo” y 4 “Muy de acuerdo”.

6 Métodos de estimación

6.1 Intención de Tratamiento (ITT)

Se evaluaron los efectos del programa sobre dos resultados primarios, dos resultados secundarios y tres resultados intermedios. Todos los resultados se expresaron en términos de tamaño del efecto en desviaciones estándar¹ y se evaluaron con niveles de confianza del 90 %, 95 % y 99 %.

El análisis ITT se realizó utilizando regresiones lineales donde las variables dependientes fueron los puntajes de las escalas generales y subescalas de: PICCOLO, ASQ-3 E2P y PSI-4-SF. La ecuación 1 representa el modelo utilizado para medir el impacto del programa sobre las distintas variables de resultado

$$Y_i = \beta_1 + \beta_2 C J_i + \lambda C i c l o_i + \gamma C o m u n a_i + \beta_3 E_{0ik} + \beta_4 X_{ni} + \mu_i \quad (1)$$

Donde Y_i es un outcome de interés, $C J_i$ es una variable dummy que toma el valor 1 si es que la diada i fue asignada al grupo de tratamiento y β_2 captura el efecto promedio del programa sobre quienes fueron asignados al tratamiento. Por su parte, E_{0ik} representa los puntajes de la encuesta inicial en la dimensión k .

Adicionalmente, $C i c l o$ es una variable categórica que indica el trimestre del año en el que fue medida la diada. Esta variable permite controlar por factores no observados que pueden variar sistemáticamente según el momento en que los individuos responden los instrumentos de evaluación. El primer ciclo se compone por quienes responden el instrumento entre los meses de mayo y julio de 2024, el segundo por las diadas que responden entre agosto y octubre de 2024, y el último grupo por aquellas que respondieron entre noviembre de 2024 y los primeros días de febrero del año 2025.

Por su parte $C o m u n a$ es una variable categórica que controla por las diferencias existentes entre los distintos territorios donde se aplica la intervención.

Por último, en cuanto a la matriz X_{ni} , se cuenta con la siguiente lista de características de las diadas, que pueden ser utilizadas como variables de control, en función de posibles diferencias que se pudiesen encontrar entre los grupos tratamiento y control:

- Nivel educacional cuidador/a
- Género cuidador/a
- Género niño/a
- Edad del cuidador/a
- Edad del niño/a
- Nacionalidad
- Relación diada (padre, madre, tío/a, etc.)
- ¿Niño/a asiste a establecimiento educacional?

¹ Los tamaños del efecto fueron estandarizados utilizando la desviación estándar del grupo de control, siguiendo la recomendación metodológica de la Education Endowment Foundation (EEF, 2022).

- ¿Niño/a tiene alguna discapacidad?
- Situación laboral cuidador/a
- Número del hijo/a (1ro, 2do, 3ro, etc.)

6.2 Límite Local de Tratamiento Promedio (LATE)

El análisis LATE se centró en las díadas que efectivamente participaron en el tratamiento, definido como aquellas que cumplieron con los criterios de participación establecidos en detalle en la sección 9.2. Para esto, se utilizó un procedimiento de variables instrumentales en dos etapas (2SLS), donde en una primera etapa, la asignación aleatoria se ocupa como variable instrumental para estimar la participación en el programa, y en una segunda etapa la variable de participación insesgada se utilizó para estimar el efecto del programa. De esta forma se estiman las siguientes ecuaciones:

$$T_i = \alpha_1 + \alpha_2 C J_i + \alpha_3 E_{0i} + \alpha_4 X_{ni} + \epsilon_i \quad (2)$$

Donde T_i es una variable que toma el valor 1 si es que la díada i fue efectivamente tratada y cero de lo contrario. A partir de los valores estimados de la ecuación 2, se realiza la estimación de la siguiente ecuación:

$$Y_i = \beta_1 + \beta_2^{LATE} \hat{T}_i + \lambda Ciclo_i + \gamma Comuna_i + \beta_3 E_{0ik} + \beta_4 X_{ni} + \mu_i \quad (3)$$

Donde T_i representa la probabilidad estimada de ser tratado por el programa y β_2^{LATE} corresponde al efecto sobre los *compliers*, es decir, el efecto sobre aquellas díadas que reciben el programa dado que fueron asignadas al grupo tratamiento. Dicho de otro modo, el β_2^{LATE} identifica el efecto causal que se centra en las díadas que efectivamente fueron tratadas por el programa.

Más adelante en la sección 9.2, se establecen los umbrales de asistencia e interacción mínima que definen la participación de las díadas en el programa.

6.3 Efectos heterogéneos

Dado que en algunas comunas la intervención se realizó en modalidad *online*, se analizó si existían diferencias en el impacto del tratamiento entre la modalidad presencial y *online*. Para medir esta interacción, se incluyó una variable en la ecuación de regresión que reflejaba el mayor o menor efecto del tratamiento cuando la intervención fue presencial. La ecuación correspondiente es:

$$Y_i = \beta_1 + \beta_2 C J_i + \beta_3 C J_i P_i + \beta_4 P_i + \lambda Ciclo_i + \gamma Comuna_i + \beta_5 E_{0ik} + \beta_6 X_{ni} + \mu_i \quad (4)$$

Donde P_i es una variable dicotómica que toma el valor 1 si la díada i pertenece a un ciclo donde la intervención se realizó de forma presencial y cero si fue *online*. De esta forma β_2 representa el efecto del tratamiento para el grupo *online*, y β_3 captura el efecto adicional (si es que este fuera positivo) que entregaría la intervención presencial con respecto a la *online*. Así, el efecto del tratamiento presencial sería $\beta_2 + \beta_3$.

Adicionalmente se realiza una estimación del efecto diferenciado del programa por grupo estario. Para esto se estima la siguiente ecuación:

$$Y_i = \beta_1 + \beta_2 C J_i + \alpha_j C J_i \sum_{j=1}^2 G_j + \delta_j \sum_{j=1}^2 G_j + \lambda Ciclo_i + \gamma Comuna_i + \beta_3 E_{0ik} + \beta_4 X_{ni} + \mu_i \quad (5)$$

Donde G_j es una variable dicotómica que toma el valor 1 si es que la díada fue asignada al grupo etario j . De esta forma, α_j indica el efecto adicional de los grupos exploradores y creadores con respecto al grupo acurrucados.

Tanto para el análisis de los efectos heterogéneos según el grupo etario, así como para identificar diferencias en los efectos según la modalidad de intervención, se estiman tanto el efecto ITT como LATE.

Cabe señalar que estos análisis presentan un menor poder estadístico, dado que la estimación de efectos en subgrupos reduce el tamaño muestral efectivo en cada comparación. Por lo tanto, los resultados de esta sección deben interpretarse con especial cautela, ya que es posible que efectos reales no alcancen significancia estadística debido a esta limitación.

6.4 Efectos marginales

Determinar la cantidad óptima de sesiones necesarias para generar un impacto significativo en las díadas es una cuestión central en la intervención. Actualmente, el programa contempla un total de 10 sesiones, aunque no existe evidencia concluyente respecto a si esta cantidad es suficiente, excesiva o insuficiente para alcanzar los objetivos propuestos. Hasta ahora, la definición del número de sesiones se ha basado principalmente en la experiencia acumulada por la Fundación Infancia Primero a lo largo del tiempo. No obstante, se requiere un análisis más riguroso que permita identificar el número óptimo de sesiones, con el objetivo de maximizar la eficiencia del programa y focalizar de forma más efectiva los recursos disponibles.

Para abordar esto, se realiza un análisis de los efectos marginales del programa, con el fin de identificar a partir de qué sesión se comienzan a observar impactos relevantes, y si existe un punto a partir del cual los beneficios adicionales disminuyen o se estabilizan. Este tipo de análisis permite tomar decisiones sobre una posible reducción, mantención o ampliación de la cantidad de sesiones, buscando un equilibrio entre impacto y sostenibilidad del programa.

Esto se realiza a través de una estimación en dos etapas. En la primera etapa, se evalúa el efecto de haber sido asignado al grupo de tratamiento sobre la cantidad de asistencias a las sesiones, controlando por las variables de caracterización y los puntajes de la prueba inicial. A partir de esta primera etapa, se calculan los residuos, tal como se expresa en la ecuación (6). En la segunda etapa, se realiza una regresión sobre las variables de impacto, incluyendo un polinomio con términos cuadráticos y cúbicos de la cantidad de sesiones a las que asiste la díada $f(NS_i)$

La elección de utilizar una función cúbica se explica por la necesidad de obtener efectos marginales que no sean lineales, una función cúbica es apropiada para estudiar efectos marginales que se comportan de forma más flexible, posibilitando que la representación gráfica del impacto se vea como una curva. Esto es importante para entender cómo cambia el beneficio adicional de la asistencia a cada sesión, reflejando mejor la realidad de que el efecto podría ser creciente a niveles bajos de asistencia, estabilizarse y luego decrecer o comportarse de forma más compleja en niveles altos. Adicionalmente, se agrega un polinomio cúbico de los residuos estimados en la primera etapa, junto con las variables de caracterización.

$$NS_i = \beta_1 + \beta_2 T_i + \beta_3 X_i + \mu_i \quad (6)$$

$$Y_i = \alpha_1 + \alpha_2 f(NS_i) + \alpha_3 X_i + f(\hat{\mu}_i) + \delta_i \quad (7)$$

6.5 Costo-efectividad

Una cuestión fundamental en la evaluación del programa es su escalabilidad, entendida como la capacidad de ser implementado en un mayor número de comunas sin que ello comprometa su efectividad.

Para abordar esta dimensión, se realiza un análisis de costos asociado a la intervención, con el objetivo de calcular un ratio de costo-efectividad. Este se define como el cociente entre el costo unitario por beneficiario y el impacto del programa, medido en desviaciones estándar. Este indicador no solo proporciona una medida del rendimiento económico del programa, sino que también permite compararlo con otras intervenciones similares, facilitando así la toma de decisiones respecto a su posible expansión.

Por lo tanto, los pasos para determinar el nivel de costo-efectividad son los siguientes (Dhaliwal, Duflo, Glennerster & Tulloch, C. 2020):

1. **Determinar el costo unitario de la intervención:** A partir de la información proporcionada por la fundación, se identifican todos los recursos utilizados durante los cinco años de implementación y se calcula el costo total, dividiéndolo por el número de díadas atendidas en ese período.
2. **Calcular el índice de costo-efectividad:** Una vez estimado el costo unitario, se utilizan los efectos del programa (expresados en desviaciones estándar) para calcular el costo asociado a lograr una mejora de una desviación estándar en los indicadores de resultado del programa.
3. **Comparar los índices de costo-efectividad:** Finalmente, se contrastan los resultados obtenidos con los de otras evaluaciones de programas de naturaleza similar al programa Crecer Jugando, con el fin de determinar cuán costo-efectivo es en relación con intervenciones comparables.

6.6 Robustez

Se realizaron pruebas de robustez con el objetivo de identificar posibles fuentes de sesgo, así como su dirección, y así evaluar la solidez y estabilidad de los resultados obtenidos.

- **Análisis de atrición:** Se efectuó un análisis de diferencia de medias entre los grupos que completaron la investigación y aquellos que desertaron. Esto con el objetivo de corroborar si la pérdida de muestra es aleatoria o esta correlacionada con alguna característica particular de las díadas.
- **Corrección por sesgo de selección:** Como análisis de robustez frente a posibles sesgos por atrición no aleatoria, se implementa un modelo de corrección por selección tipo Heckman (Heckman, 1979). Este modelo permite estimar el efecto del programa ajustando por el hecho de que la muestra observada podría estar sesgada debido a diferencias sistemáticas entre quienes permanecen en el estudio y quienes no.

Finalmente, tras cada estimación, se aplicó la corrección de Romano-Wolf para ajustar por pruebas múltiples, considerando la correlación entre los distintos resultados analizados.²

²La corrección de Romano-Wolf controla la tasa de error familiar (FWER, por sus siglas en inglés). Esta corrección es más poderosa que procedimientos como las correcciones de Bonferroni y Holm, ya que toma en cuenta la estructura de dependencia de las estadísticas de

7 Trabajo de campo

El levantamiento de información se llevó a cabo en 6 comunas del país a lo largo del año 2024, periodo en el que se inició la implementación de hasta tres ciclos y 2 subgrupos por comunas. Si bien todos los ciclos de implementación del programa comenzaron en 2024, la recolección de información se extendió hasta febrero de 2025.

7.1 Procedimiento de aleatorización y contacto

Al comienzo de cada periodo, se aplicó una entrevista y encuesta de entrada tanto a los grupos de control como de tratamiento, y una vez finalizada la intervención, se aplicaron distintos instrumentos para medir la efectividad de la intervención en diversas variables de interés. Para contar con mejores tasas de respuesta, se definió la entrega de una gift card de \$20.000 (a canjear en diferentes tiendas nacionales) a los participantes del grupo de control y tratamiento que completaran las actividades consideradas para efectos de este estudio.

En cada comuna y ciclo, sobre el universo de postulantes que se inscribieron en la plataforma, se asignó aleatoriamente el 50 % al grupo control y el otro 50 % al grupo tratamiento. Para cada grupo, se siguió el siguiente protocolo de contacto:

Grupo Tratamiento:

1. A partir de la semana 9 del tratamiento, se contactó telefónicamente a las díadas para invitarlas a participar del estudio y agendar la aplicación del instrumento PICCOLO. Las díadas estaban al tanto de que serían contactadas por las investigadoras gracias a la información proporcionada por los equipos implementadores.
2. Finalizado el llamado telefónico, se envió a través de WhatsApp el enlace del cuestionario I (E2P y PSI-4) y el consentimiento informado para participar en la evaluación.
3. Se aplicó el instrumento PICCOLO en la fecha y horario convenido. Una vez finalizada su aplicación, se envió el enlace del cuestionario II (ASQ-3 y encuesta de salida) a través de WhatsApp.
4. Una vez desarrolladas todas las actividades, se le hace envío de la gift card a la participante a través de WhatsApp.

Grupo Control:

1. Se contactó telefónicamente a los participantes para informarles que quedaron en la lista de espera para recibir el tratamiento y se les indicó la fecha aproximada de inicio de su participación.
2. Se les invitó a participar de la evaluación y, a quienes accedieron, se les aplicó la entrevista y encuesta inicial.
3. Una vez contactadas todas las díadas, se envió un correo electrónico con el consentimiento informado para participar en la evaluación.
4. A partir de la semana 7 de la intervención del grupo tratamiento, se contactó telefónicamente a este grupo para aplicar el cuestionario I (E2P y PSI-4) y agendar el instrumento PICCOLO. Se envió el enlace del cuestionario I a través de WhatsApp para su autoaplicación.

prueba mediante remuestreo de los datos originales.

5. Se aplicó el instrumento PICCOLO en la fecha y horario convenido y, al finalizar su aplicación, se envió el enlace del cuestionario II (ASQ-3 y encuesta de salida) vía WhatsApp.
6. Una vez desarrolladas todas las actividades, se le hace envío de la gift card a la participante a través de WhatsApp.

7.2 Nivel de respuesta y ajustes por atrición

En la Tabla 1 se observa el plan inicial de aleatorización. En una primera instancia, a partir de información provista por Infancia Primero, se proyectó una atrición del 15 %, por lo que, para llegar a 460 díadas, era necesario aleatorizar 540 en total.

Tabla 1: Plan inicial de aleatorización

Comuna	Muestra	Ciclo 1		Ciclo 2		Ciclo 3	
		T	C	T	C	T	C
Concepcion	150	39	39	12	12	24	24
El Bosque	140	21	21	27	27	22	22
Limache	36	0	0	10	10	8	8
Maule	46	0	0	15	15	8	8
Melipilla	54	15	15	12	12	0	0
Monte Patria	40	9	9	11	11	0	0
Quinta Normal	52	10	10	8	8	8	8
Rio Negro	24	12	12	0	0	0	0
Total	542	106	106	95	95	70	70

Fuente: elaboración propia.

Tabla 2: Aleatorización por comuna

Comuna	Muestra	Ciclo 1		Ciclo 2		Ciclo 3	
		T	C	T	C	T	C
Concepcion	218	68	71	39	40	0	0
El Bosque	152	22	24	27	27	26	26
Limache	31	0	0	10	10	5	6
Maule	61	15	16	0	0	15	15
Melipilla	47	15	16	0	0	7	9
Monte Patria	27	0	0	14	11	1	1
Quinta Normal	42	11	11	3	4	7	6
Rio Negro	22	12	10	0	0	0	0
Total	600	143	148	93	92	61	63

Fuente: elaboración propia.

A lo largo de la investigación se identificaron dos momentos en los que se producía una atrición de la muestra participante. En un primer momento, luego de la aleatorización inicial, existe un grupo de cuidadores/as que no responde la encuesta y entrevista inicial. A este primer desistimiento lo hemos denominado Atrición Etapa 1. Las razones identificadas son 3: En primer lugar, existen cuidadores/as, parte del grupo control, que no desean participar de la investigación y no responden los instrumentos iniciales (entrevistas y encuestas). Un segundo grupo está compuesto por cuidadores/as que al ser contactados declaran que, si bien se inscribieron en algún momento para recibir el programa, por razones personales, ya no quieren participar (ya sea del grupo control o tratamiento). Por último, está el grupo de cuidadores/as a los que no fue posible contactar.

Un segundo momento de pérdida muestral se identifica durante la aplicación de los instrumentos finales del estudio (PICCOLO, ASQ-3, E2P y PSI-4), lo que denominamos Atrición Etapa 2. Esta etapa de atrición obedece principalmente a dos factores. En primer lugar, la negativa de algunos participantes a ser grabados para la aplicación del instrumento PICCOLO, lo que constituye una forma de abandono parcial del estudio, ya que en varios casos estos participantes sí completan los otros instrumentos. En segundo lugar, se observa un abandono completo asociado a la falta de respuesta a los instrumentos, incluso tras reiterados intentos de contacto.

La Tabla 2 muestra el total de díadas aleatorizadas por comuna. Como puede observarse, el número total de díadas aleatorizadas tuvo que ser incrementado respecto al plan inicial debido a los niveles de atrición observados. Este ajuste permitió contar con una base más amplia que contribuyera a aumentar la cantidad de participantes efectivos en la evaluación.

En la Tabla 3 se presenta la atrición de la Etapa 1 por comuna. En primer lugar, se observan niveles de atrición disímiles tanto entre las comunas, como también entre los grupos tratamiento y control. Por ejemplo, en el grupo tratamiento, la comuna de Concepción presenta niveles de atrición del 27 %, valor muy distinto a la comuna de Melipilla que alcanza una tasa de respuesta del 100 %. Estas diferencias son atribuibles a variables de contexto y gestiones realizadas por los equipos territoriales.

Por su parte, en el grupo control se observa que, por ejemplo, la comuna de El Bosque tiene un nivel de atrición que alcanza el 19 %, mientras que las comunas de Limache, Maule y Melipilla presentan una atrición del 7 % y Río Negro del 0 %. Estas diferencias de atrición entre la comuna de El Bosque y el resto de los territorios, probablemente se deba a que esta última fue la comuna pionera y piloto para el protocolo de contacto, el cual fue ajustado luego de identificar una serie de oportunidades de mejora.

En cuanto a la atrición en la Etapa 2, en la Tabla 4 se observan porcentajes totales de respuesta similares en ambos grupos, con algunas diferencias en la atrición de las distintas comunas.

Durante el desarrollo del trabajo de campo, el equipo de investigación fue monitoreando activamente las razones de desistimiento por parte de los/as cuidadores/as contactados/as, entre las que destacan: desinformación sobre el programa y su evaluación, especialmente en participantes del grupo control o con baja vinculación en el grupo tratamiento; resistencia al uso de instrumentos grabados; dificultades logísticas o personales (como falta de tiempo o cambio de número telefónico); y casos en que no fue posible establecer contacto pese a múltiples intentos. Estas situaciones dieron lugar a la implementación progresiva de estrategias para mejorar la tasa de respuesta, como ajustes en el discurso, reforzamiento del vínculo territorial y búsqueda alternativa de contacto.

Cabe mencionar que en la Tabla 4 se presentan aquellas díadas que participan de la investigación, es decir que responden alguno de los cuatro instrumentos (PICCOLO, ASQ-3, E2P y PSI-4). Más adelante, en la sección de resultados, se considerará como muestra a las díadas que hayan respondido el instrumento correspondiente al análisis.

Tabla 3: Atrición etapa 1

	Entrevista completada			Atrición Etapa 1
	No	Sí	Total	
Tratamiento	62	241	303	20 %
Concepcion	30	81	111	27 %
El Bosque	10	67	77	13 %
Limache	4	12	16	25 %
Maule	3	28	31	10 %
Melipilla	0	25	25	0 %
Monte Patria	6	6	12	50 %
Quinta Normal	6	15	21	29 %
Rio Negro	3	7	10	30 %
Control	35	262	297	12 %
Concepcion	12	95	107	11 %
El Bosque	14	61	75	19 %
Limache	1	14	15	7 %
Maule	2	28	30	7 %
Melipilla	2	20	22	9 %
Monte Patria	1	14	15	7 %
Quinta Normal	3	18	21	14 %
Rio Negro	0	12	12	0 %
Total	97	503	600	16 %

Fuente: elaboración propia.

Nuevamente, dado que los niveles de atrición en ambos grupos supera el 30 %, es necesario analizar si esta pérdida de muestra podría generar sesgos en la estimación del efecto del programa. Más adelante, en la sección de Análisis de Robustez, se realizarán algunos ejercicios para identificar en qué medida la atrición podría generar un sesgo en la estimación..

En la Tabla 5 se observa la cantidad de respuestas por instrumento y finalmente en la Tabla 6 un resumen de las etapas de atrición. En esta última se evidencia que en una primera Etapa se produce una pérdida de 97 díadas seguido de una pérdida de 179, lo que da como resultado un total de 276 díadas que se pierden en el proceso, es decir, un porcentaje de atrición del 46 %.

Tabla 4: Atrición etapa 2

	Participa en la investigación			Atrición Etapa 2
	No	Sí	Total	
Tratamiento	87	154	241	36 %
Concepcion	33	48	81	41 %
El Bosque	25	42	67	37 %
Limache	1	11	12	8 %
Maule	12	16	28	43 %
Melipilla	9	16	25	36 %
Monte Patria	0	6	6	0 %
Quinta Normal	5	10	15	33 %
Rio Negro	2	5	7	29 %
Control	92	170	262	35 %
Concepcion	40	55	95	42 %
El Bosque	22	39	61	36 %
Limache	3	11	14	21 %
Maule	7	21	28	25 %
Melipilla	6	14	20	30 %
Monte Patria	5	9	14	36 %
Quinta Normal	2	16	18	11 %
Rio Negro	7	5	12	58 %
Total	179	324	503	36 %

Fuente: elaboración propia.

Tabla 5: Respuesta instrumentos

	Entrevista	Encuesta inicial	Encuesta final	PSI4	E2P	ASQ3	Picc.
Control	262	261	193	163	166	162	97
Tratamiento	241	247	208	162	161	149	74
Total	503	508	401	325	327	311	171

Fuente: elaboración propia

Tabla 6: Resumen atrición

Aleatorizadas	Completan la entrevista	Participan de la investigación
	503 (84 %)	324 (54 %)
	Atrición Etapa 1	Atrición Etapa 2
600 (100 %)	97 (16 %)	179 (36 %)
	Atrición total	
	276 (46 %)	

Fuente: elaboración propia.

8 Cálculos Efecto Mínimo Detectable

El cálculo del efecto mínimo detectable (MDES por sus siglas en inglés) está dado por la siguiente ecuación propuesta por Duflo, Glennerster y Kremer (2006):

$$MDES = t_{\alpha} + t_{1-\kappa} \sqrt{\sigma^2 / P(1 - P)N} \quad (8)$$

Donde t_{α} y $t_{1-\kappa}$ son los estadísticos t evaluados en el nivel de significancia α y poder estadístico $1 - \kappa$, respectivamente. La constante P indica la proporción de la muestra que es tratada, N el tamaño de la muestra.

Dado que los individuos son seleccionados forma aleatoria, podemos asumir que las observaciones son i.i.d (variables independientes e idénticamente distribuidas) con varianza σ^2 .

Para evaluar el efecto mínimo detectable, se consideró un nivel de significancia del 5 %, un poder estadístico del 80 % y una proporción de tratados y controles de 50/50. Estos valores son un estándar en estudios de impacto que buscan minimizar el error Tipo 1 (significancia del 5 %) y garantizar una alta probabilidad de detectar un efecto verdadero (poder del 80 %). En la Tabla 7 se muestra el efecto mínimo detectable calculado para el total de respuestas de cada uno de los instrumentos.

Se puede observar que inicialmente se calculó una muestra de 540 (270 tratamientos y 270 control), que luego de proyectar una atrición de 15 %, permitía obtener un MDES de 0.23 DE. Posteriormente, se observa que la atrición fue superior a la esperada, con un MDES de 0.28 DE para todos los instrumentos, a excepción del PICCOLO donde el tamaño de la muestra permitiría detectar un efecto mínimo de 0.30 DE.

Estudios sobre intervenciones similares, como el realizado en EE. UU. sobre el programa PC-CARE, una intervención de 7 sesiones cuyo objetivo es mejorar la calidad de la relación entre cuidadores y niños de 2 a 10 años, mediante apoyo y entrenamiento en habilidades parentales, encontraron una reducción del estrés parental superior a 0.3 DE (Hawk et al., 2022). Asimismo, Jeong et al. (2022) llevaron a cabo un metanálisis de 111 ensayos controlados aleatorios de intervenciones con menores de 3 años destinadas a fortalecer el vínculo cuidador-infantes. Los autores encontraron tamaños de efecto de 0.56 DE en conocimientos sobre crianza, 0.33 DE en prácticas de crianza y 0.39 DE en interacción. A partir de esta evidencia, podemos concluir que los MDE calculados para esta investigación podrían ser suficientemente bajos.

Tabla 7: Efecto Mínimo Detectable

	Protocolo	PSI-4-SF	E2P	ASQ-3	PICC.
MDES	0.23	0.28	0.28	0.28	0.30
Medidas	460	325	327	311	281
Tratamiento	230	162	161	149	136
Control	230	163	166	162	145
Aleatorizadas	540	600	600	600	600
Tratamiento	270	297	297	297	297
Control	270	303	303	303	303
Atrición	15 %	46 %	46 %	48 %	59 %
Tratamiento	15 %	45 %	46 %	50 %	63 %
Control	15 %	46 %	45 %	47 %	56 %

Fuente: elaboración propia.

9 Resultados

9.1 Caracterización de los participantes

Para observar las diferencias existentes entre los grupos tratamiento y control se realiza una caracterización de las diadas que completaron su proceso de investigación. Es decir, diadas que respondieron al menos uno de los 4 instrumentos mencionados anteriormente (E2P, ASQ3, PSI4 o PICCOLO). En la Tabla 8 se observa una caracterización en base a la entrevista inicial.

Se evidencia, en primer lugar, una alta proporción de cuidadoras de género femenino, con un 99 % en el grupo Tratamiento y un 97 % en el grupo Control. En cuanto a las diferencias entre los/as cuidadores/as de ambos grupos, se identifican diferencias estadísticamente significativas en dos características: la proporción de nacionalidad chilena y la proporción con empleo. En el primer caso, el grupo de Tratamiento presenta una mayor proporción de cuidadores/as de nacionalidad chilena, mientras que en el segundo, el grupo de Control exhibe una mayor proporción de cuidadores/as con empleo.

En relación a las diferencia en las características de los infantes entre ambos grupos, se observan diferencias estadísticamente significativas en la proporción que son primeros hijo/as, con una mayor proporción en el grupo control.

En cuanto a las diferencias en la encuesta inicial, en la Tabla 9 se observa el balance en la frecuencia de interacción, el apoyo social percibido por los/as cuidadores/as y las habilidades de buen trato. Nuevamente en este caso se observan diferencias estadísticamente significativas entre ambos grupos, siendo el grupo tratamiento el que comienza en una posición mas desaventajada. Es decir, el grupo que recibe el programa tienen una menor frecuencia de interacción, menores habilidades de buen trato y menor apoyo social percibido que el grupo control.

Para evitar posibles sesgos provenientes de este desbalance, en las estimaciones de los efectos se incorporarán los puntajes iniciales como variables de control.

Tabla 9: Balance encuesta inicial

	Control	Tratamiento	Dif.	Std. Error	N
Frecuencia de interacciones	18.79	17.46	1.34**	(0.48)	315
Apoyo social percibido	8.89	8.35	0.53*	(0.26)	315
Habilidades de buen trato	15.83	15.28	0.55*	(0.30)	315
Dificultades parentales	4.68	4.94	-0.26	(0.35)	315

Fuente: elaboración propia.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Tabla 8: Caracterización general díadas

	Control	Trat.	Dif.	Std. Error	N
Cuidadores/as					
Proporción de mujeres cuidadoras	0.97	0.99	-0.02	(0.02)	322
Edad promedio cuidador/a	32.36	31.52	0.84	(0.79)	321
Proporción que son madres del infante	0.95	0.98	-0.03	(0.02)	322
Proporción de cuidadores/as de nacionalidad chilena	0.85	0.91	-0.06*	(0.04)	322
Proporción de cuidadores/as con educación superior completa o incompleta	0.63	0.65	-0.02	(0.05)	322
Proporción de cuidadores/as casados o con pareja conviviente	0.57	0.61	-0.03	(0.06)	321
Proporción de cuidadores/as con empleo formal o informal	0.59	0.44	0.15**	(0.06)	322
Infantes					
Proporción infantes de género femenino	0.55	0.49	0.06	(0.05)	338
Edad promedio de los infantes	17.99	17.50	0.50	(1.39)	338
Proporción de infantes que asisten a algún establecimiento educacional	0.36	0.31	0.05	(0.05)	322
Proporción que es primer hijo/a	0.78	0.67	0.11*	(0.05)	322
Proporción de infantes con alguna discapacidad	0.08	0.04	0.04	(0.03)	322

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nota: Las diferencias en el N en la proporción de infantes de género femenino y su edad promedio, se deben a que esta información es asignada directamente al momento de la inscripción de los/as participantes.

9.2 Fidelidad de la implementación

En las comunas que se incorporan en el presente reporte, la implementación del programa se realizó dentro del tiempo planificado, iniciando entre marzo y noviembre de 2024 y terminando en enero del año 2025. En cada una de ellas se llevaron a cabo el total de 10 sesiones proyectadas, que considera a su vez la implementación de la sesión (presencial u *online* según sea el caso) y la provisión de recursos, herramientas y contenidos asociados vía Whatsapp (cuento, audio, canción y guías).

En la Tabla 10, se presenta la distribución temporal de las 340 díadas participantes en la evaluación (control y tratamiento), desagregadas según el mes de inicio de implementación del ciclo y comuna. Es relevante señalar que las comunas de Concepción y El Bosque se encuentran subdivididas en dos grupos. En el caso de Concepción, la diferenciación responde a la modalidad de la intervención, distinguiéndose entre participantes en formato presencial y *online*. Por su parte, en El Bosque, si bien ambos subgrupos corresponden a intervenciones presenciales, estas se llevan a cabo en distintas instalaciones de la comuna. Del total de díadas participantes de la evaluación, un 24,4 % postuló a una modalidad *online* y otro 75,6 % a una presencial.

Tabla 10: Distribución de díadas según mes de inicio de implementación de ciclo, por comuna

	Mar	Abr	May	Jul	Ago	Sep	Oct	Nov
Concepción (Online)	0	42	0	0	18	0	0	0
Concepción	0	26	0	0	23	0	0	0
El Bosque 1	16	0	0	16	0	0	16	0
El Bosque 2	0	0	0	20	0	0	15	0
Limache	0	0	0	17	0	0	6	0
Maule	0	0	18	0	0	20	0	0
Melipilla	0	21	0	0	0	9	0	0
Monte Patria (Online)	0	0	0	0	17	0	0	2
Quinta Normal (Online)	0	14	0	6	0	0	0	7
Rio Negro	0	0	11	0	0	0	0	0

Fuente: elaboración propia.

Para identificar la participación efectiva de las díadas participantes de la evaluación y asignadas al grupo tratamiento, se utilizó la información recolectada durante el proceso de monitoreo. Estos datos permiten evaluar la exposición real de los participantes a las intervenciones. Las variables se agrupan en dos categorías principales: participación por asistencia a sesiones y participación a través de grupo de Whatsapp.

9.2.1. Participación por asistencia a sesiones

La participación en las sesiones se evalúa a través de la asistencia a estas durante un ciclo de intervención que consta de 10 sesiones, realizadas semanalmente, tanto en formato online como presencial. En conjunto con el equipo técnico de Infancia Primero, se estableció que se considerará participación efectiva cuando un participante asista a un mínimo de 4 sesiones.

En la Tabla 11 se muestra la distribución de la asistencia en el grupo tratamiento, evidenciándose que el 52,5 % de los participantes alcanzó la participación efectiva, es decir, asistieron a 4 o más sesiones.

Tabla 11: Asistencia a sesiones del grupo tratamiento (presencial u online)

Asistencia	N	%	Cum. %
Sin asistencia	38	22.6	22.6
1	19	11.3	33.9
2	12	7.1	41.1
3	11	6.5	47.6
4	15	8.9	56.5
5	11	6.5	63.1
6	10	6.0	69.0
7	14	8.3	77.4
8	22	13.1	90.5
9	8	4.8	95.2
10	8	4.8	100.0
Total	168	100	

Fuente: elaboración propia.

9.2.2. Participación vía WhatsApp

La participación a través de WhatsApp corresponde a la interacción de los/as participantes en un grupo donde reciben información adicional relacionada con el contenido de cada sesión, y donde tienen la posibilidad de interactuar y discutir sobre temas relacionados a la intervención. La participación se registró mediante la descarga de audios, canciones, y conversaciones vinculadas a los contenidos de las sesiones durante todo el ciclo. Para definir la participación efectiva en WhatsApp, se consideró que los/as participantes debían haber descargado audios, canciones o participado activamente en las 10 sesiones del ciclo.

De acuerdo con los registros del grupo de tratamiento, en la Tabla 12 se observa que el 28,6 % de los participantes alcanzó la participación efectiva vía WhatsApp.

9.2.3. Participación efectiva en el programa

Finalmente, se desarrolló un indicador global de participación, que integró ambas categorías: asistencia a sesiones y participación vía grupo de WhatsApp. Se consideró que un participante tuvo participación efectiva si cumplió con al menos uno de los dos criterios establecidos. De acuerdo con los datos, el 63,7% de los participantes cumplió con alguno de los criterios de participación.

9.3 Resultados de impacto

A continuación, se presentan los resultados intermedios, primarios y secundarios. En todos los casos, se reportan los resultados para tres modelos distintos, empleando dos tipos de estimadores: Intention-to-Treat (ITT) y Local Average Treatment Effect (LATE).

Tabla 12: Participación vía WhatsApp del grupo tratamiento

Participación	N	%	Cum. %
Sin participación	2	1.2	1.2
2	4	2.4	3.6
3	3	1.8	5.4
4	8	4.8	10.1
5	10	6.0	16.1
6	16	9.5	25.6
7	12	7.1	32.7
8	25	14.9	47.6
9	40	23.8	71.4
10	48	28.6	100.0
Total	168	100	

Fuente: elaboración propia.

Tabla 13: Participación efectiva del grupo tratamiento

	N	%
Participa	107	63.7
No participa	61	36.3
Total	168	100 %

Fuente: elaboración propia.

El primer modelo incluye únicamente la variable de tratamiento, junto a controles por el trimestre en que se midió a la diada y a la comuna donde se realizó la intervención. El segundo modelo, además de estas variables, controla por los puntajes obtenidos en la encuesta inicial en aquellas dimensiones donde se identificaron diferencias significativas entre los grupos de tratamiento y control (frecuencia de interacción, habilidades de buen trato y apoyo social percibido). Por último, un tercer modelo controla por las características de las diadas donde se observaron diferencias significativas entre los grupo.

En relación al cálculo de los tamaños de efecto, los coeficientes están estandarizados por la desviación estándar del grupo control. Por su parte, con respecto a las estimaciones por 2SLS, la evaluación de la primera etapa muestra un estadístico F superior a 100 en todos los modelos estimados, lo que indica que el instrumento tiene una fuerte capacidad predictiva (Staiger y Stock, 1997).

9.3.1. Resultados intermedios

En cuanto a los resultados intermedios, al controlar por los puntajes iniciales y las características de las diadas, en la Tabla 14 se observa un efecto positivo y significativo de 0.24 DE en la frecuencia de interacción al considerar el estimador LATE.

En relación con las habilidades parentales, tanto el estimador ITT como el LATE presentan efectos positivos y sig-

nificativos en los tres modelos para todas las dimensiones, con la excepción de las habilidades vinculares. En esta última dimensión, el efecto solo alcanza significancia estadística una vez que se controla por los puntajes iniciales y las características de las díadas. En cuanto a la magnitud del efecto, se observa un impacto particularmente relevante en las habilidades reflexivas, con un tamaño de efecto de 0.39 DE.

Todos los coeficientes LATE de los resultados intermedios pasan la prueba de múltiples hipótesis usando la metodología Romano-Wolf.

Tabla 14: Resultados intermedios

	ITT			LATE		
	(1)	(2)	(3)	(4)	(5)	(6)
Frecuencia de interacción	0.065 (0.091)	0.183 (0.109)	0.159 (0.090)	0.100 (0.126)	0.280* (0.151)	0.240** (0.118)
Habilidades Parentales						
Habilidades vinculares	0.114 (0.102)	0.213* (0.093)	0.198* (0.088)	0.175 (0.152)	0.327** (0.146)	0.299** (0.137)
Habilidades formativas	0.153* (0.080)	0.271** (0.074)	0.240** (0.089)	0.235* (0.128)	0.415** (0.128)	0.363** (0.146)
Habilidades protectoras	0.137* (0.061)	0.226*** (0.042)	0.203** (0.059)	0.209** (0.085)	0.347*** (0.060)	0.307*** (0.088)
Habilidades Reflexivas	0.170* (0.088)	0.266** (0.097)	0.255* (0.121)	0.260** (0.117)	0.408** (0.126)	0.386** (0.163)
Control por puntajes iniciales	No	Sí	Sí	No	Sí	Sí
Control por características	No	No	Sí	No	No	Sí

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

9.3.2. Resultados primarios

En relación con los resultados primarios, la Tabla 15 muestra un efecto estadísticamente significativo de 0.24 DE en el desarrollo socio-individual de los infantes. En el resto de las dimensiones del ASQ3, si bien se observan efectos positivos, estos no alcanzan significancia estadística.

En cuanto a la interacción parental, se observa un efecto estadísticamente significativo del programa en la dimensión enseñanza, con un tamaño de efecto de 0.31 DE.

9.3.3. Resultados secundarios

La Tabla 16 muestra los resultados del programa en el estrés parental y el apoyo social percibido. En cuanto al primero, se observa que el programa tiene un efecto negativo y significativo en todas las dimensiones, con especial énfasis en la angustia parental, donde se observa una reducción de 0.39 DE. Con respecto al estrés total, se evidencia que el programa genera una caída de 0.41 DE.

Por otro lado, en cuanto al apoyo social percibido, no se observan efectos significativos en ninguna de las especificaciones del modelo

Tabla 15: Resultados primarios

	ITT			LATE		
	(1)	(2)	(3)	(4)	(5)	(6)
Desarrollo Infante - ASQ3						
Comunicación	0.051 (0.100)	0.053 (0.106)	0.054 (0.111)	0.079 (0.148)	0.084 (0.156)	0.084 (0.159)
Motora fina	0.077 (0.090)	0.100 (0.093)	0.108 (0.091)	0.121 (0.134)	0.158 (0.142)	0.167 (0.140)
Motora gruesa	0.022 (0.163)	0.014 (0.160)	0.017 (0.155)	0.035 (0.234)	0.021 (0.231)	0.027 (0.220)
Resolución de problemas	0.036 (0.079)	0.029 (0.110)	0.038 (0.103)	0.056 (0.114)	0.046 (0.158)	0.059 (0.145)
Socio-individual	0.118* (0.059)	0.135 (0.075)	0.161* (0.076)	0.184* (0.094)	0.212* (0.115)	0.249** (0.114)
Interacción Parental - Piccolo						
Afecto	-0.010 (0.099)	-0.004 (0.099)	-0.017 (0.104)	-0.016 (0.138)	-0.006 (0.137)	-0.026 (0.142)
Responsividad	-0.031 (0.087)	0.008 (0.094)	0.002 (0.110)	-0.046 (0.121)	0.011 (0.130)	0.004 (0.150)
Aliento	0.057 (0.099)	0.082 (0.116)	0.060 (0.100)	0.086 (0.136)	0.124 (0.158)	0.091 (0.136)
Enseñanza	0.200** (0.079)	0.211** (0.078)	0.207* (0.090)	0.303** (0.138)	0.320** (0.135)	0.312** (0.148)
Total	0.077 (0.041)	0.104 (0.056)	0.090 (0.058)	0.117** (0.059)	0.158** (0.080)	0.135 (0.084)
Control por puntajes iniciales	No	Sí	Sí	No	Sí	Sí
Control por características	No	No	Sí	No	No	Sí

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Finalmente, todos los coeficientes LATE de los resultados secundarios pasan la prueba de múltiples hipótesis usando la metodología Romano-Wolf.

Tabla 16: Resultados secundarios

	ITT			LATE		
	(1)	(2)	(3)	(4)	(5)	(6)
Estrés Parental - PSI4						
Angustia Parental	-0.161*	-0.259**	-0.257**	-0.248**	-0.396***	-0.387***
	(0.082)	(0.051)	(0.060)	(0.114)	(0.072)	(0.068)
Interacción disfuncional	-0.130	-0.200*	-0.202*	-0.200	-0.306**	-0.304**
	(0.109)	(0.103)	(0.095)	(0.156)	(0.149)	(0.134)
Características difíciles	-0.152*	-0.220**	-0.219**	-0.233**	-0.336**	-0.329**
	(0.071)	(0.074)	(0.083)	(0.101)	(0.105)	(0.116)
Estrés total	-0.179*	-0.275**	-0.275**	-0.275**	-0.421***	-0.413***
	(0.087)	(0.070)	(0.068)	(0.123)	(0.102)	(0.092)
Apoyo social percibido	0.011	0.087	0.063	0.017	0.133	0.095
	(0.084)	(0.086)	(0.089)	(0.118)	(0.120)	(0.121)
Control por puntajes iniciales	No	Sí	Sí	No	Sí	Sí
Control por características	No	No	Sí	No	No	Sí

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

9.4 Efectos marginales

Para estimar el impacto en la intensidad del tratamiento, es necesario aplicar un modelo que permita identificar la existencia de efectos no lineales para cada sesión adicional a la que asista una díada. Es decir, se explora si el efecto de asistir a una sesión más estaría condicionado por la cantidad de sesiones a las que ya se ha asistido.

La Tabla 17 muestra que la asistencia a las primeras sesiones no presenta efectos marginales significativos sobre las habilidades parentales. Sin embargo, a partir de la tercera o cuarta sesión, comienzan a observarse efectos positivos y estadísticamente significativos en todas las dimensiones evaluadas.

De manera similar, en cuanto al desarrollo infantil, la Tabla 18 indica que en el área de habilidades socioindividuales se detecta un efecto marginal positivo a partir de la tercera sesión. No obstante, a partir de la séptima sesión, el efecto marginal de una sesión adicional deja de ser significativo, lo que sugiere la existencia de un umbral a partir del cual el impacto marginal del programa desaparece.

En las dimensiones de comunicación y motricidad gruesa, se observan efectos marginales negativos en los primeros niveles de asistencia (0 o 1 sesión). Este patrón podría reflejar un fenómeno de autoselección negativa, en el que las familias con mayores dificultades o menor compromiso inicial tienden a abandonar tempranamente. Alternativamente, las primeras sesiones podrían generar una mayor conciencia por parte del cuidador respecto a posibles retrasos del niño, lo que se traduce en evaluaciones más críticas. En conjunto, estos factores explican la dinámica inicial negativa, seguida por una reversión del efecto: a partir de la quinta sesión, se observan impactos positivos sostenidos en motricidad gruesa.

Por su parte, en la Tabla 19 al analizar los efectos marginales en la dimensión Enseñanza (única en la que se observaron efectos estadísticamente significativos) se observa que los efectos aparecen únicamente cuando las díadas han asistido a dos sesiones del programa.

Por último, en la Tabla 20 al evaluar los resultados secundarios, se observa que una vez que las díadas han asistido a 3 sesiones, el efecto marginal del programa es negativo en todas las dimensiones del estrés parental. Sin embargo, estos efectos desaparecen una vez que las díadas han asistido a más cinco sesiones.

Los resultados del análisis de efectos marginales sugieren que el programa Crecer Jugando comienza a generar impactos positivos en las habilidades parentales, el estrés parental y las habilidades socioindividuales a partir de la asistencia a al menos tres sesiones. Asimismo, se observa que, una vez superado el umbral de cinco sesiones (siete en el caso del desarrollo de habilidades socioindividuales), el efecto marginal de asistir a una sesión adicional deja de ser estadísticamente significativo.

Sin embargo, este último hallazgo debe interpretarse con cautela, dado que el número de díadas que asiste a más de siete sesiones es reducido, lo que limita la precisión de las estimaciones en ese rango.

En consecuencia, los resultados entregan indicios de que la asistencia continua al programa es relevante a partir de la tercera sesión, y plantean como una línea futura de investigación el estudio más detallado del valor agregado de asistir a más de siete sesiones.

En las Figuras 3,4 y 5 y 6 se observa gráficamente el efecto marginal para cada una de las dimensiones de los distintos resultados.

Tabla 17: Habilidades parentales, E2P (Efecto Marginal)

	(1) Vinculares	(2) Formativas	(3) Protectoras	(4) Reflexivas
0	-0.129 (0.101)	-0.043 (0.069)	-0.058 (0.120)	-0.117 (0.079)
1	-0.035 (0.053)	0.013 (0.031)	0.003 (0.062)	-0.014 (0.048)
2	0.042 (0.026)	0.057* (0.028)	0.047 (0.032)	0.063 (0.035)
3	0.100** (0.039)	0.088* (0.045)	0.072 (0.047)	0.116** (0.034)
4	0.141* (0.060)	0.107* (0.056)	0.079 (0.066)	0.142*** (0.034)
5	0.163* (0.077)	0.113 (0.060)	0.069 (0.077)	0.144*** (0.033)
6	0.167 (0.092)	0.107 (0.057)	0.040 (0.079)	0.121** (0.040)
7	0.153 (0.109)	0.088 (0.053)	-0.007 (0.077)	0.072 (0.064)
8	0.121 (0.132)	0.057 (0.061)	-0.072 (0.083)	-0.002 (0.104)
9	0.071 (0.166)	0.013 (0.091)	-0.155 (0.109)	-0.102 (0.158)
10	0.003 (0.212)	-0.043 (0.138)	-0.256 (0.159)	-0.227 (0.224)
N	327	327	327	326

Fuente: elaboración propia. Nota: Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla 18: Desarrollo infante, ASQ3 (Efecto Marginal)

	(1) Comunicación	(2) Motora Fina	(3) Motora Gruesa	(4) Resolución Problemas	(5) Socio Individual
0	-0.167*** (0.029)	0.011 (0.167)	-0.370** (0.120)	-0.079 (0.170)	-0.340 (0.184)
1	-0.091*** (0.022)	0.021 (0.084)	-0.193** (0.059)	-0.046 (0.094)	-0.148 (0.087)
2	-0.023 (0.030)	0.030 (0.027)	-0.045 (0.049)	-0.015 (0.033)	0.005 (0.028)
3	0.040 (0.045)	0.038 (0.057)	0.071 (0.075)	0.012 (0.042)	0.119* (0.053)
4	0.095 (0.062)	0.045 (0.099)	0.157 (0.097)	0.037 (0.087)	0.194** (0.081)
5	0.145 (0.080)	0.051 (0.131)	0.212* (0.108)	0.059 (0.127)	0.230** (0.092)
6	0.187 (0.099)	0.055 (0.152)	0.237* (0.111)	0.079 (0.160)	0.226** (0.090)
7	0.223 (0.120)	0.059 (0.163)	0.232* (0.116)	0.095 (0.189)	0.184* (0.092)
8	0.252 (0.143)	0.061 (0.169)	0.196 (0.135)	0.109 (0.214)	0.102 (0.129)
9	0.275 (0.167)	0.062 (0.173)	0.129 (0.180)	0.120 (0.239)	-0.019 (0.205)
10	0.291 (0.194)	0.062 (0.184)	0.032 (0.251)	0.129 (0.268)	-0.179 (0.314)
N	311	311	311	311	311

Fuente: elaboración propia. Nota: Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla 19: Vínculo Parental, PICCOLO (Efecto Marginal)

	(1) Afecto	(2) Aliento	(3) Responsividad	(4) Enseñanza	(5) Total
0	0.204 (0.190)	0.014 (0.160)	-0.004 (0.188)	0.185 (0.237)	0.117 (0.225)
1	0.096 (0.102)	-0.003 (0.079)	-0.010 (0.103)	0.115 (0.111)	0.058 (0.113)
2	0.003 (0.034)	-0.007 (0.023)	-0.017 (0.035)	0.061** (0.022)	0.014 (0.025)
3	-0.073* (0.038)	0.003 (0.055)	-0.027 (0.038)	0.023 (0.076)	-0.017 (0.059)
4	-0.132 (0.079)	0.025 (0.098)	-0.038 (0.084)	0.001 (0.132)	-0.034 (0.114)
5	-0.175 (0.110)	0.061 (0.133)	-0.050 (0.121)	-0.004 (0.168)	-0.037 (0.151)
6	-0.201 (0.134)	0.109 (0.160)	-0.065 (0.148)	0.006 (0.188)	-0.026 (0.171)
7	-0.211 (0.156)	0.171 (0.183)	-0.082 (0.167)	0.033 (0.197)	-0.001 (0.176)
8	-0.204 (0.184)	0.246 (0.205)	-0.100 (0.179)	0.076 (0.209)	0.037 (0.170)
9	-0.181 (0.226)	0.334 (0.233)	-0.120 (0.189)	0.135 (0.240)	0.090 (0.162)
10	-0.142 (0.286)	0.436 (0.272)	-0.142 (0.201)	0.210 (0.303)	0.157 (0.170)
N	281	281	281	281	281

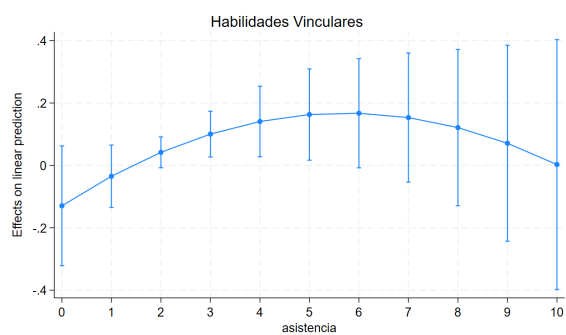
Fuente: elaboración propia. Nota: Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla 20: Estrés parental, PSI-4 (Efecto Marginal)

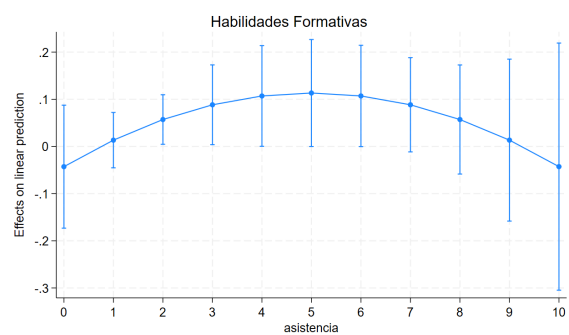
	(1) Angustia Parental	(2) Interacción disfuncional	(3) Características Difíciles	(4) Estrés Total
0	0.159 (0.163)	0.079 (0.163)	0.040 (0.154)	0.117 (0.154)
1	0.043 (0.082)	0.002 (0.082)	-0.029 (0.078)	0.010 (0.078)
2	-0.049 (0.034)	-0.060* (0.034)	-0.077** (0.033)	-0.073** (0.033)
3	-0.116** (0.056)	-0.108* (0.056)	-0.107** (0.053)	-0.132** (0.053)
4	-0.159* (0.089)	-0.141 (0.089)	-0.116 (0.084)	-0.167** (0.084)
5	-0.178 (0.115)	-0.158 (0.115)	-0.107 (0.108)	-0.178 (0.108)
6	-0.173 (0.135)	-0.161 (0.135)	-0.077 (0.127)	-0.165 (0.127)
7	-0.143 (0.157)	-0.150 (0.157)	-0.029 (0.148)	-0.127 (0.148)
8	-0.090 (0.190)	-0.123 (0.190)	0.040 (0.180)	-0.066 (0.180)
9	-0.012 (0.241)	-0.081 (0.241)	0.128 (0.228)	0.019 (0.228)
10	0.090 (0.314)	-0.025 (0.314)	0.235 (0.297)	0.129 (0.297)
N	325	325	325	325

Fuente: elaboración propia. Nota: Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

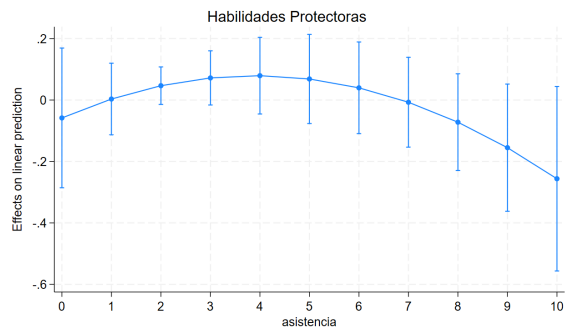
Figura 3: Efecto Marginal Habilidades Parentales



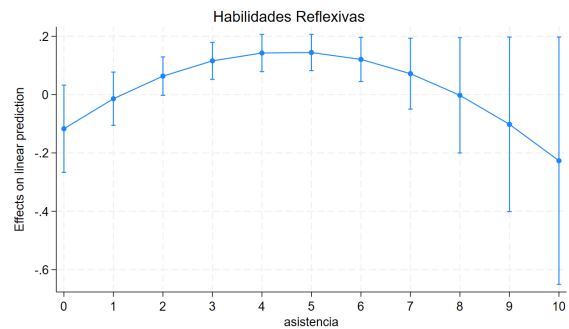
(a)



(b)



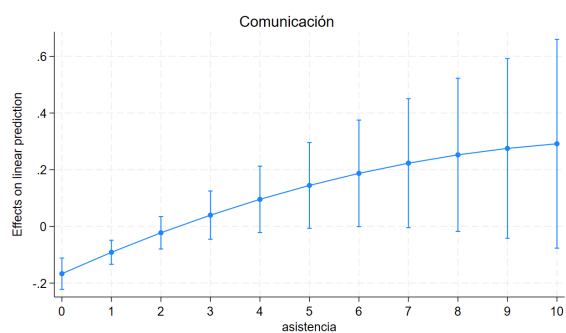
(c)



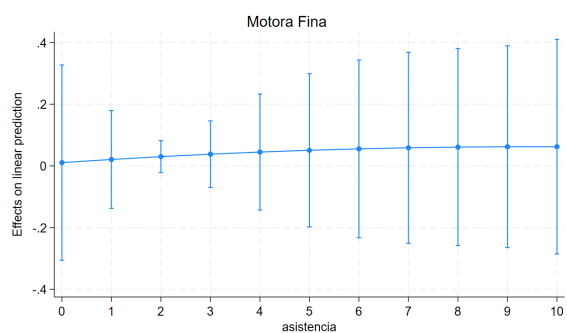
(d)

Fuente: elaboración propia. Las líneas verticales representan un intervalo de confianza al 90 % de confianza.

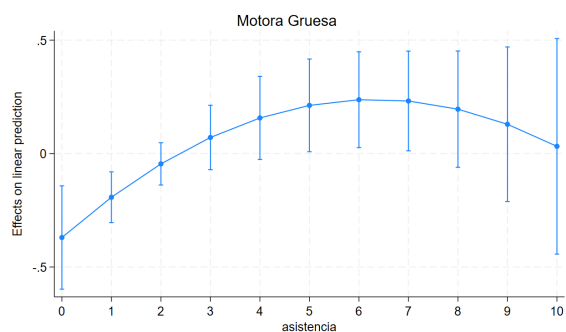
Figura 4: Efecto Marginal Desarrollo del Infante



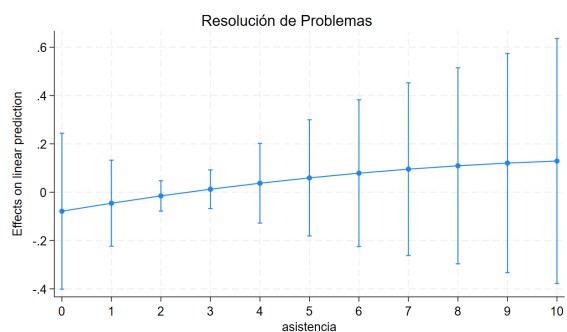
(a)



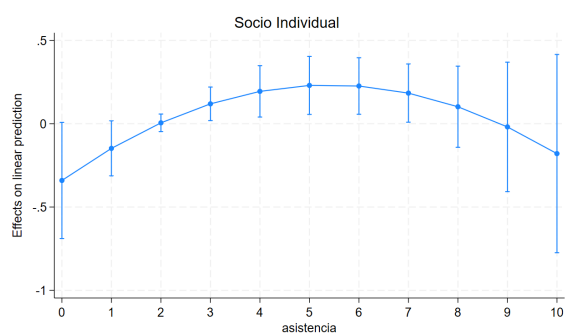
(b)



(c)



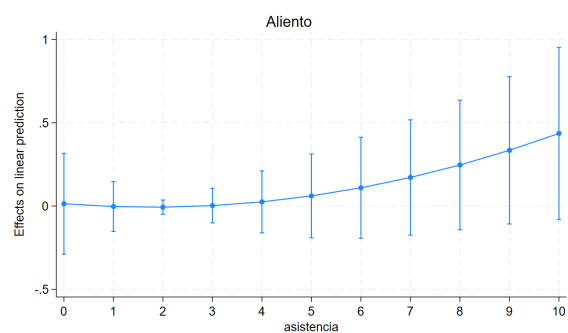
(d)



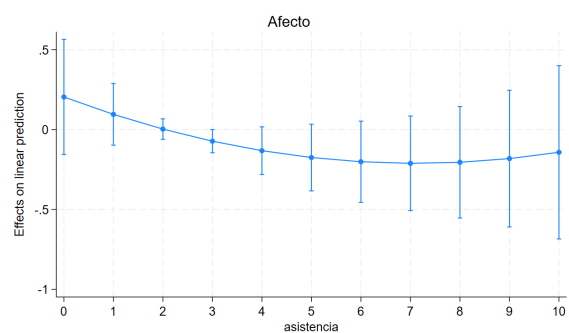
(e)

Fuente: elaboración propia. Las líneas verticales representan un intervalo de confianza al 90 % de confianza.

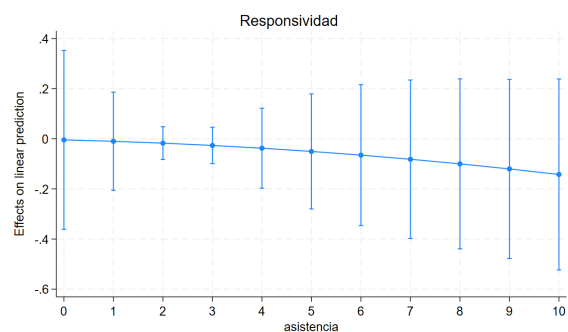
Figura 5: Efecto Marginal Interacción Parental



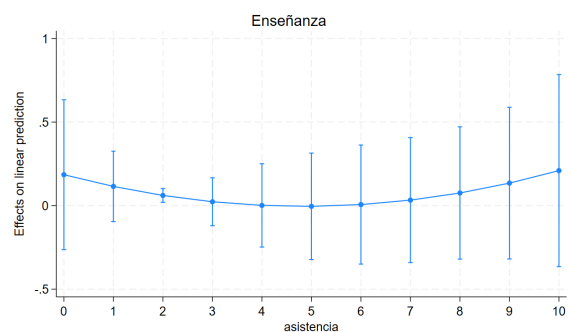
(a)



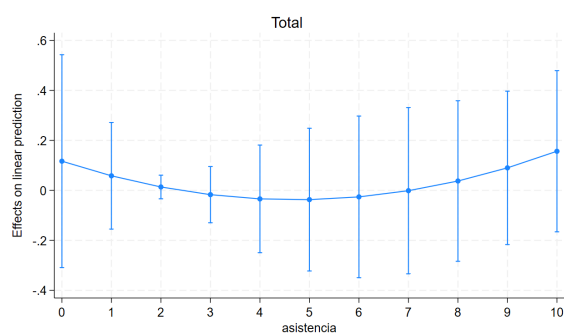
(b)



(c)



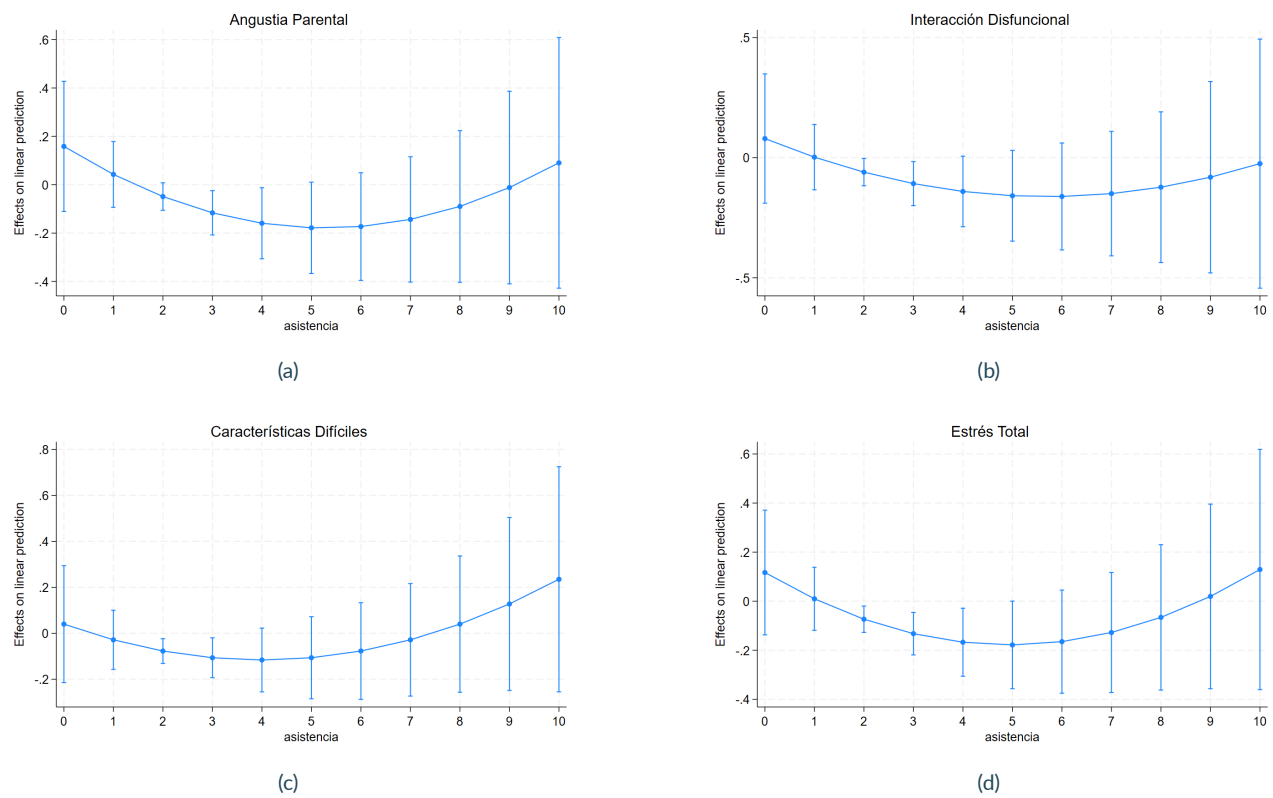
(d)



(e)

Fuente: elaboración propia. Las líneas verticales representan un intervalo de confianza al 90 % de confianza.

Figura 6: Efecto Marginal Estrés Parental



Fuente: elaboración propia. Las líneas verticales representan un intervalo de confianza al 90 % de confianza.

9.5 Efectos heterogéneos

9.5.1. Efectos heterogéneos sobre modalidad de tratamiento

Una de las características particulares del programa es que su implementación adopta dos modalidades distintas: presencial y *online*. Esta diferenciación responde a decisiones adoptadas por cada comuna, en función de las particularidades de sus contextos, criterios de pertinencia y condiciones de factibilidad operativa.

Dado lo anterior, resulta relevante indagar si existen diferencias en los efectos del programa según la modalidad de implementación, reconociendo que estas diferencias pueden estar influenciadas tanto por la modalidad misma como por las características de los territorios que optaron por una u otra forma de ejecución.

Esta diferencia se puede analizar mediante un efecto de interacción entre la variable de tratamiento y una variable que indica si la díada recibió el tratamiento de manera presencial u *online*.

En la Tabla 21 se presenta una comparación entre los grupos control y tratamiento dentro del subconjunto de díadas que se inscribieron voluntariamente en la modalidad presencial del programa. Se identifican diferencias estadísticamente significativas en el nivel educativo y la situación laboral. Por su parte en cuanto a las características de los infantes, se observan diferencias en el porcentaje que corresponde al primer hijo/a y a la proporción que tiene alguna discapacidad.

Por otro lado, la Tabla 22 realiza la misma comparación entre tratamiento y control, pero esta vez solo para el grupo *online*. Al realizar esta comparación solo se observa una diferencia estadísticamente significativa en el porcentaje de cuidadores/as que tiene educación superior.

Para solucionar en parte el posible sesgo producido por el desbalance entre los grupos tanto en las díadas que reciben el tratamiento *online* como presencial, se agregan como controles a la regresión aquellas variables donde se observan diferencias en cada uno de los grupos.

Es importante señalar que al evaluar el efecto del tratamiento según subgrupos de modalidad (presencial versus *online*), se produce una reducción en la potencia estadística de las estimaciones. Esto se debe a que al dividir la muestra total en subgrupos más pequeños, se reduce el tamaño efectivo de muestra disponible para cada análisis, lo que puede limitar la capacidad para detectar efectos estadísticamente significativos. Adicionalmente, el grupo que recibió el tratamiento en modalidad *online* está compuesto por 33 díadas, mientras que el grupo presencial considera 135 díadas, lo que da cuenta de una asimetría importante en el tamaño de los subgrupos.

Dicho lo anterior, en las columna (2) y (5) de la Tablas 23 se observa que la modalidad presencial tiene un efecto superior a la modalidad *online* en las habilidades protectoras, sin embargo, al controlar por las características de las díadas, la significancia estadística desaparece.

Por su parte, en cuanto a los resultados primarios, en la Tabla 24 se observa que el grupo que recibe el tratamiento presencial tiene mayores efectos en comunicación y desarrollo de motricidad fina por el lado de las dimensiones del ASQ3. Por otro lado, con respecto a la interacción parental, al controlar por los puntajes iniciales y por las características de las díadas, se observa que no existe una diferencia en los efectos entre la modalidad *online* y presencial.

Finalmente, en relación con los resultados secundarios, la Tabla 25 muestra que únicamente el indicador de apoyo social percibido presenta efectos significativamente mayores en la modalidad presencial. Este hallazgo es completamente coherente con lo esperado, dado que dicha modalidad permite una mayor interacción entre los cuidadores, lo que podría favorecer la percepción de apoyo social.

Tabla 21: Caracterización díadas modalidad presencial

	Control	Trat.	Dif.	Std. Error	N
Cuidadores/as					
Proporción de mujeres cuidadoras	0.97	0.98	-0.02	(0.02)	248
Edad promedio cuidador/a	32.04	30.88	1.16	(0.92)	248
Proporción que son madres del infante	0.94	0.98	-0.03	(0.02)	248
Proporción de cuidadores/as de nacionalidad chilena	0.88	0.93	-0.05	(0.04)	248
Proporción de cuidadores/as con educación superior completa o incompleta	0.54	0.66	-0.11*	(0.06)	248
Proporción de cuidadores/as casados o con pareja conviviente	0.62	0.62	-0.01	(0.06)	248
Proporción de cuidadores/as con empleo formal o informal	0.57	0.41	0.16*	(0.06)	248
Infantes					
Proporción infantes de género femenino	0.52	0.48	0.04	(0.06)	257
Edad promedio de los infantes	16.95	17.39	-0.43	(1.56)	257
Proporción de infantes que asisten a algún establecimiento educacional	0.33	0.26	0.07	(0.06)	248
Proporción que es primer hijo/a	0.76	0.64	0.12*	(0.06)	248
Proporción de infantes con alguna discapacidad	0.08	0.03	0.05*	(0.03)	248

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nota: Las diferencias en el N en la proporción de infantes de género femenino y su edad promedio, se deben a que esta información es asignada directamente al momento de la inscripción de los/as participantes.

Tabla 22: Caracterización diadas modalidad online

	Control	Trat.	Dif.	Std. Error	N
Cuidadores/as					
Proporción de mujeres cuidadoras	0.98	1.00	-0.02	(0.03)	76
Edad promedio cuidador/a	33.12	35.12	-2.00	(1.54)	75
Proporción que son madres del infante	0.98	1.00	-0.02	(0.03)	76
Proporción de cuidadores/as de nacionalidad chilena	0.76	0.77	-0.01	(0.10)	76
Proporción de cuidadores/as con educación superior completa o incompleta	0.82	0.62	0.20*	(0.10)	76
Proporción de cuidadores/as casados o con pareja conviviente	0.48	0.48	0.00	(0.12)	75
Proporción de cuidadores/as con empleo formal o informal	0.64	0.58	0.06	(0.12)	76
Infantes					
Proporción infantes de género femenino	0.62	0.48	0.14	(0.11)	83
Edad promedio de los infantes	20.28	17.67	2.61	(3.05)	83
Proporción de infantes que asisten a algún establecimiento educacional	0.44	0.62	-0.18	(0.12)	76
Proporción que es primer hijo/a	0.80	0.81	-0.01	(0.10)	76
Proporción de infantes con alguna discapacidad	0.08	0.08	0.00	(0.07)	76

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nota: Las diferencias en el N en la proporción de infantes de género femenino y su edad promedio, se deben a que esta información es asignada directamente al momento de la inscripción de los/as participantes.

Tabla 23: Efectos heterogéneos de modalidad de tratamiento resultados intermedios

	ITT			LATE		
	(1)	(2)	(3)	(4)	(5)	(6)
Frecuencia de interacción	-0.015 (0.195)	0.105 (0.149)	-0.058 (0.135)	-0.015 (0.254)	0.179 (0.184)	-0.065 (0.172)
Habilidades Parentales - E2P						
Habilidades vinculares	0.112 (0.160)	0.169 (0.139)	0.070 (0.161)	0.180 (0.223)	0.278 (0.197)	0.139 (0.234)
Habilidades formativas	0.175 (0.140)	0.253 (0.153)	0.140 (0.191)	0.280 (0.223)	0.409 (0.252)	0.253 (0.292)
Habilidades protectoras	0.148 (0.092)	0.145* (0.075)	0.081 (0.079)	0.237* (0.125)	0.242** (0.088)	0.159 (0.104)
Habilidades Reflexivas	-0.033 (0.142)	-0.017 (0.138)	-0.084 (0.145)	-0.029 (0.194)	0.007 (0.183)	-0.078 (0.200)
Control por puntajes iniciales	No	Sí	Sí	No	Sí	Sí
Control por características	No	No	Sí	No	No	Sí

Fuente: elaboración propia. Nota: Los coeficientes indican el efecto adicional de la modalidad presencial sobre la modalidad online. Entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla 24: Efectos heterogéneos de modalidad de tratamiento resultados primarios

	ITT			LATE		
	(1)	(2)	(3)	(4)	(5)	(6)
Desarrollo Infante - ASQ3						
Comunicación	0.566** (0.135)	0.585** (0.144)	0.578** (0.243)	0.855*** (0.219)	0.882*** (0.231)	0.866** (0.382)
Motora fina	0.291** (0.105)	0.329** (0.108)	0.221* (0.111)	0.447** (0.167)	0.509** (0.174)	0.340* (0.191)
Motora gruesa	0.036 (0.392)	0.044 (0.384)	0.021 (0.298)	0.055 (0.519)	0.066 (0.503)	0.018 (0.380)
Resolución de problemas	-0.080 (0.291)	-0.061 (0.289)	-0.116 (0.194)	-0.115 (0.405)	-0.088 (0.395)	-0.181 (0.260)
Socio-individual	0.151 (0.123)	0.156 (0.125)	0.089 (0.093)	0.242 (0.183)	0.252 (0.187)	0.151 (0.154)
Interacción Parental - Piccolo						
Afecto	0.259 (0.151)	0.296 (0.170)	0.169 (0.264)	0.381* (0.200)	0.432* (0.222)	0.240 (0.340)
Responsividad	-0.091 (0.318)	-0.087 (0.313)	-0.201 (0.363)	-0.139 (0.425)	-0.130 (0.413)	-0.307 (0.469)
Aliento	0.229 (0.193)	0.266 (0.221)	0.057 (0.323)	0.340 (0.262)	0.395 (0.297)	0.087 (0.428)
Enseñanza	0.147 (0.195)	0.181 (0.179)	0.055 (0.199)	0.228 (0.285)	0.282 (0.263)	0.108 (0.296)
Total	0.171 (0.135)	0.207 (0.154)	0.019 (0.250)	0.256 (0.185)	0.310 (0.207)	0.034 (0.329)
Control por puntajes inicial	No	Sí	Sí	No	Sí	Sí
Control por características	No	No	Sí	No	No	Sí

Fuente: elaboración propia. Nota: Los coeficientes indican el efecto adicional de la modalidad presencial sobre la modalidad online. Entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla 25: Efectos heterogéneos de modalidad de tratamiento resultados secundarios

	ITT			LATE		
	(1)	(2)	(3)	(4)	(5)	(6)
Estrés Parental - PSI4						
Angustia Parental	-0.022 (0.299)	-0.041 (0.319)	-0.029 (0.336)	-0.041 (0.404)	-0.073 (0.425)	-0.068 (0.445)
Interacción disfuncional	0.158 (0.408)	0.132 (0.424)	0.344 (0.393)	0.235 (0.556)	0.194 (0.570)	0.516 (0.521)
Características difíciles	0.073 (0.340)	0.070 (0.367)	0.204 (0.341)	0.104 (0.460)	0.099 (0.492)	0.296 (0.457)
Estrés total	0.069 (0.402)	0.050 (0.429)	0.177 (0.412)	0.097 (0.544)	0.066 (0.573)	0.250 (0.547)
Apoyo social percibido	0.446** (0.132)	0.379** (0.151)	0.281* (0.143)	0.630*** (0.185)	0.551** (0.201)	0.401* (0.208)
Control por puntajes iniciales	No	Sí	Sí	No	Sí	Sí
Control por características	No	No	Sí	No	No	Sí

Fuente: elaboración propia. Nota: Los coeficientes indican el efecto adicional de la modalidad presencial sobre la modalidad online. Entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

9.5.2. Efectos heterogéneos sobre grupo etario asignado

Otra de las características del programa es que la intervención se ajusta según la edad del infante. Según los meses de edad, la díada es asignada a uno de tres grupos: Acurrucados, Exploradores o Creadores. Ante esta asignación diferenciada, surge la pregunta de si existe algún efecto heterogéneo del tratamiento según el grupo al que pertenece la díada.

Para investigar este efecto heterogéneo, se agregan a la regresión principal dos variables de interacción: una entre la variable de tratamiento y una *dummy* que indica si la díada pertenece al grupo Creadores, y otra entre la variable de tratamiento y una *dummy* que indica si la díada pertenece al grupo Exploradores. En este contexto, los resultados deben interpretarse como el efecto adicional del tratamiento sobre el grupo base, que en este caso son aquellas díadas asignadas al grupo Acurrucados.

Tabla 26: Diferencias entre grupo control y tratamiento según grupo etario

	Acurrucados	Exploradores	Creadores
Cuidadores/as			
Proporción de mujeres cuidadoras	-0.01	-0.05	0.01
Edad promedio cuidador/a	-0.62	2.02	1.54
Proporción que son madres del infante	0.00	-0.02	-0.07
Proporción de cuidadores/as de nacionalidad chilena	-0.02	-0.11	-0.06
Proporción de cuidadores/as con educación superior completa o incompleta	-0.09	0.06	0.01
Proporción de cuidadores/as casados o con pareja conviviente	-0.06	0.02	-0.02
Proporción de cuidadores/as con empleo formal o informal	0.18*	0.14	0.16
Infantes			
Proporción infantes de género femenino	0.06	-0.02	0.18*
Edad promedio de los infantes	-0.27	-0.55	-1.92
Proporción de infantes que asisten a algún establecimiento educacional	-0.05	0.08	0.07
Proporción que es primer hijo/a	-0.02	0.08	0.37**
Proporción de infantes con alguna discapacidad	0.03	0.04	0.07
N	161	88	91

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas. Los valores representan las diferencias entre el grupo control y el grupo tratamiento para cada grupo etario. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Para evaluar el balance de características entre los grupos, la Tabla 26 presenta las diferencias observadas en cada uno de ellos. Se identifican diferencias en la proporción de género femenino y en la proporción de casos en que se trata del primer hijo en el grupo Creadores, así como una diferencia estadísticamente significativa en la

proporción de cuidadoras con empleo en el grupo Acurrucados. Dado lo anterior, estas variables se incorporan como controles en el modelo para ajustar por las diferencias de base en cada uno de los grupos.

Dicho lo anterior, en la Tabla 27 se observa que no existen diferencias entre los grupos estarios en la frecuencia de interacción. Sin embargo, se evidencia que, en el grupo exploradores, el programa genera un mayor impacto, tanto en habilidades protectoras como reflexivas.

Por otro lado, en las Tablas 28 y 29 se observa que el programa impacta con mayor intensidad la comunicación en el grupo exploradores, con respecto al grupo acurrucados. En la dimensión resolución de problemas se observa que el grupo creadores tiene un menor efecto que el grupo acurrucados, sin embargo, esta diferencia desaparece una vez que se controla por las características de las díadas.

En relación al vínculo, se observa que el programa impacta en menor medida al grupo explorador con respecto a el grupo acurrucados tanto en las dimensiones enseñanza como responsividad.

Por último, en la Tabla 30 se observa que el programa reduce en mayor medida la angustia parental y el estrés total en los grupos exploradores y creadores. De igual forma en la dimensión características difíciles se observa una mayor reducción en el grupo creadores. En el apoyo social percibido no se observan diferencias.

Tabla 27: Efectos heterogéneos de grupos etarios (resultados intermedios)

		ITT			LATE		
		(1)	(2)	(3)	(4)	(5)	(6)
Frecuencia de interacción							
Exploradores		0.194 (0.211)	0.107 (0.222)	0.113 (0.207)	0.214 (0.306)	0.134 (0.317)	0.129 (0.294)
Creadores		-0.166 (0.371)	-0.315 (0.274)	-0.326 (0.206)	-0.357 (0.668)	-0.500 (0.515)	-0.559 (0.384)
Habilidades Parentales - E2P							
Habilidades vinculares							
Exploradores		0.188 (0.246)	0.061 (0.252)	0.075 (0.248)	0.261 (0.321)	0.083 (0.332)	0.101 (0.319)
Creadores		0.299 (0.250)	0.184 (0.218)	0.172 (0.234)	0.557 (0.449)	0.424 (0.385)	0.371 (0.401)
Habilidades formativas							
Exploradores		0.327 (0.278)	0.180 (0.298)	0.197 (0.304)	0.461 (0.364)	0.255 (0.394)	0.277 (0.389)
Creadores		0.376 (0.223)	0.242 (0.157)	0.235 (0.201)	0.695 (0.457)	0.539 (0.342)	0.481 (0.407)
Habilidades protectoras							
Exploradores		0.595** (0.190)	0.461* (0.206)	0.460* (0.205)	0.850*** (0.221)	0.658** (0.235)	0.654** (0.226)
Creadores		0.470 (0.387)	0.363 (0.317)	0.302 (0.378)	0.810 (0.668)	0.683 (0.552)	0.526 (0.646)
Habilidades Reflexivas							
Exploradores		0.755** (0.203)	0.617** (0.229)	0.626** (0.239)	1.106*** (0.178)	0.898*** (0.226)	0.910*** (0.232)
Creadores		0.089 (0.391)	-0.030 (0.308)	-0.063 (0.367)	0.134 (0.656)	-0.008 (0.524)	-0.112 (0.625)
Control por puntajes inicial		No	Sí	Sí	No	Sí	Sí
Control por características		No	No	Sí	No	No	Sí

Fuente: elaboración propia. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nota: Los coeficientes indican el efecto adicional de la modalidad presencial sobre la modalidad online

Tabla 28: Efectos heterogéneos de grupos etarios (resultados primarios - desarrollo del infante)

		ITT			LATE		
		(1)	(2)	(3)	(4)	(5)	(6)
Desarrollo Infante - ASQ3							
Comunicación							
Exploradores		0.709**	0.691**	0.684**	1.043**	0.999**	0.982**
		(0.238)	(0.229)	(0.219)	(0.345)	(0.319)	(0.303)
Creadores		-0.091	-0.096	-0.036	-0.283	-0.283	-0.165
		(0.311)	(0.305)	(0.308)	(0.553)	(0.530)	(0.537)
Motora fina							
Exploradores		0.162	0.166	0.141	0.237	0.238	0.202
		(0.241)	(0.255)	(0.260)	(0.325)	(0.333)	(0.334)
Creadores		-0.025	-0.018	-0.034	-0.052	-0.030	-0.058
		(0.246)	(0.224)	(0.222)	(0.497)	(0.448)	(0.438)
Motora gruesa							
Exploradores		-0.247	-0.221	-0.214	-0.351	-0.313	-0.303
		(0.389)	(0.414)	(0.419)	(0.528)	(0.553)	(0.555)
Creadores		-0.230	-0.213	-0.170	-0.363	-0.342	-0.241
		(0.290)	(0.276)	(0.286)	(0.403)	(0.367)	(0.367)
Resolución de problemas							
Exploradores		0.238	0.237	0.228	0.363	0.350	0.333
		(0.220)	(0.209)	(0.217)	(0.305)	(0.278)	(0.285)
Creadores		-0.267	-0.255	-0.183	-0.546*	-0.524*	-0.366
		(0.163)	(0.161)	(0.133)	(0.325)	(0.315)	(0.243)
Socio-individual							
Exploradores		0.041	0.017	0.018	0.062	0.025	0.023
		(0.215)	(0.227)	(0.244)	(0.304)	(0.315)	(0.335)
Creadores		-0.101	-0.123	-0.054	-0.119	-0.142	0.020
		(0.270)	(0.256)	(0.294)	(0.493)	(0.457)	(0.529)
Control por puntajes inicial		No	Sí	Sí	No	Sí	Sí
Control por características		No	No	Sí	No	No	Sí

Fuente: elaboración propia. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nota: Los coeficientes indican el efecto adicional de los grupos exploradores y creadores sobre el grupo acurrucados.

Tabla 29: Efectos heterogéneos de grupos etarios (resultados primarios - vínculo parental)

		ITT			LATE		
		(1)	(2)	(3)	(4)	(5)	(6)
Vínculo Parental - PICCOLO							
Afecto							
	Exploradores	0.121 (0.289)	0.153 (0.299)	0.119 (0.317)	0.170 (0.360)	0.209 (0.362)	0.162 (0.380)
	Creadores	-0.044 (0.249)	-0.050 (0.252)	0.011 (0.287)	-0.108 (0.415)	-0.120 (0.424)	-0.006 (0.475)
Responsividad							
	Exploradores	-0.399 (0.296)	-0.426 (0.279)	-0.453 (0.290)	-0.546 (0.375)	-0.589* (0.356)	-0.622* (0.370)
	Creadores	-0.226 (0.228)	-0.271 (0.211)	-0.217 (0.217)	-0.367 (0.340)	-0.422 (0.310)	-0.301 (0.322)
Aliento							
	Exploradores	-0.339 (0.298)	-0.337 (0.299)	-0.373 (0.318)	-0.478 (0.381)	-0.473 (0.386)	-0.516 (0.407)
	Creadores	-0.039 (0.246)	-0.063 (0.239)	-0.026 (0.262)	0.019 (0.436)	-0.013 (0.428)	0.059 (0.462)
Enseñanza							
	Exploradores	-0.651** (0.187)	-0.615** (0.189)	-0.642** (0.199)	-0.936*** (0.241)	-0.865*** (0.259)	-0.896*** (0.263)
	Creadores	0.176 (0.245)	0.158 (0.212)	0.224 (0.254)	0.551 (0.468)	0.514 (0.411)	0.651 (0.463)
Total							
	Exploradores	-0.442 (0.279)	-0.430 (0.289)	-0.470 (0.309)	-0.625* (0.360)	-0.603 (0.379)	-0.651 (0.401)
	Creadores	-0.035 (0.253)	-0.066 (0.228)	0.002 (0.261)	0.054 (0.456)	0.010 (0.415)	0.149 (0.465)
Control por puntajes inicial		No	Sí	Sí	No	Sí	Sí
Control por características		No	No	Sí	No	No	Sí

Fuente: elaboración propia. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nota: Los coeficientes indican el efecto adicional de los grupos exploradores y creadores sobre el grupo acurruados.

Tabla 30: Efectos heterogéneos de grupos etarios (resultados secundarios)

		ITT			LATE		
		(1)	(2)	(3)	(4)	(5)	(6)
Estrés Parental - PSI4							
Angustia parental							
	Exploradores	-0.614** (0.188)	-0.512** (0.199)	-0.515** (0.192)	-0.862*** (0.245)	-0.723** (0.254)	-0.721** (0.239)
	Creadores	-0.553** (0.214)	-0.441** (0.183)	-0.431* (0.185)	-0.973** (0.412)	-0.839** (0.342)	-0.791** (0.338)
Interacción disfuncional							
	Exploradores	-0.371 (0.283)	-0.298 (0.260)	-0.284 (0.250)	-0.520 (0.379)	-0.421 (0.346)	-0.399 (0.331)
	Creadores	-0.313 (0.268)	-0.242 (0.256)	-0.200 (0.240)	-0.561 (0.486)	-0.480 (0.455)	-0.383 (0.414)
Características difíciles							
	Exploradores	-0.558 (0.391)	-0.454 (0.341)	-0.444 (0.329)	-0.786 (0.520)	-0.641 (0.447)	-0.623 (0.424)
	Creadores	-0.455** (0.178)	-0.383* (0.168)	-0.366* (0.156)	-0.792** (0.372)	-0.710** (0.324)	-0.649** (0.323)
Estrés Total							
	Exploradores	-0.634* (0.326)	-0.521 (0.302)	-0.514 (0.293)	-0.891** (0.433)	-0.735* (0.397)	-0.720* (0.379)
	Creadores	-0.546** (0.182)	-0.440** (0.141)	-0.416** (0.113)	-0.960** (0.385)	-0.837** (0.295)	-0.759** (0.260)
Apoyo social percibido							
	Exploradores	0.234 (0.232)	0.061 (0.228)	0.050 (0.228)	0.335 (0.307)	0.090 (0.307)	0.076 (0.298)
	Creadores	0.099 (0.275)	-0.032 (0.192)	-0.082 (0.203)	0.142 (0.455)	-0.021 (0.324)	-0.131 (0.344)
Control por puntajes inicial		No	Sí	Sí	No	Sí	Sí
Control por características		No	No	Sí	No	No	Sí

Fuente: elaboración propia. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nota: Los coeficientes indican el efecto adicional de los grupos exploradores y creadores sobre el grupo acurrucados.

9.6 Análisis de robustez: sesgo por atrición

Una posible fuente de sesgo en la estimación del impacto del programa es la existencia de diádas que abandonan la investigación. Esta deserción no causaría problemas de endogeneidad si ocurre de manera aleatoria y no está relacionada con las características de los cuidadores o niños evaluados. Para identificar si existe alguna característica particular en las diádas que abandonan el estudio, en la Tabla 31 se presenta un análisis de diferencia de medias entre el grupo que completó la investigación y aquel que la abandonó.

Este ejercicio revela diferencias estadísticamente significativas al 90 % de confianza en dos variables: el nivel educativo del/de la cuidador/a y la proporción de infantes que asisten a algún establecimiento educacional. En ambos casos, los valores son mayores en el grupo que completó la investigación. En el resto de las variables consideradas, no se observan diferencias estadísticamente significativas.

En relación con las dimensiones de frecuencia de interacción, apoyo social, habilidades de buen trato y dificultades parentales, la Tabla 32 muestra que no existen diferencias significativas entre quienes completaron la investigación y quienes la abandonaron. No obstante, al desagregar el análisis por grupo experimental, se identifican diferencias significativas únicamente en la frecuencia de interacción dentro del grupo de control, con valores más altos en quienes sí completaron el estudio.

Dado que una mayor frecuencia de interacción se asocia con mejores resultados parentales, la atrición podría subestimar el efecto real del programa, por lo que los resultados deben interpretarse como una estimación conservadora de su impacto.

De todas formas, como un análisis de robustez, para abordar este potencial sesgo de atrición, se implementa un modelo de corrección de selección tipo Heckman. Este modelo permite estimar el impacto del programa corrigiendo una posible no aleatoriedad en la muestra final, al modelar explícitamente el proceso de respuesta como una primera etapa (selección), e incorporar esa probabilidad en la estimación del efecto del tratamiento en la segunda etapa (resultado). La identificación del modelo se logra mediante la inclusión de una variable de exclusión válida. En este caso se utiliza un indicador del encuestador a cargo del levantamiento, que se presume afecta la probabilidad de respuesta, pero no el resultado observado.

En las Tablas 33, 34 y 35 se detecta evidencia de selección no aleatoria en la dimensión Enseñanza, dado que el coeficiente lambda resulta estadísticamente significativo. Esto indica que las personas que permanecen en la muestra y responden a esta dimensión difieren sistemáticamente, en características no observadas, de quienes no lo hacen, y que dichas diferencias están correlacionadas con el resultado. Esta selección genera un sesgo que es corregido por el modelo de Heckman, lo que explica por qué el efecto del programa deja de ser significativo tras aplicar la corrección. En contraste, en el resto de las dimensiones no se detecta evidencia de selección correlacionada con el resultado, por lo que los estimadores sin corrección pueden considerarse válidos.

Tabla 31: Caracterización general diadas según quienes completan y quienes no completan la investigación

	No completa	Sí Completa	Dif.	Std. Error	N
Cuidadores/as					
Proporción de mujeres cuidadoras	0,97	0,98	-0,01	(0,01)	503
Edad promedio cuidador/a	31,76	31,99	-0,23	(0,67)	502
Proporción que son madres del infante	0,96	0,97	-0,01	(0,02)	503
Proporción de cuidadores/as de nacionalidad chilena	0,87	0,87	0,00	(0,03)	503
Proporción de cuidadores/as con educación superior completa o incompleta	0,55	0,64	-0,08*	(0,05)	503
Proporción de cuidadores/as casados o con pareja conviviente	0,52	0,59	-0,07	(0,05)	502
Proporción de cuidadores/as con empleo formal o informal	0,45	0,52	-0,06	(0,05)	503
Infantes					
Proporción infantes de género femenino	0,49	0,52	-0,03	(0,04)	600
Edad promedio de los infantes	19,17	17,68	1,49	(1,07)	600
Proporción de infantes que asisten a algún establecimiento educacional	0,27	0,34	-0,07*	(0,04)	503
Proporción que es primer hijo/a	0,66	0,72	-0,06	(0,04)	503
Proporción de infantes con alguna discapacidad	0,06	0,06	-0,00	(0,02)	503

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla 32: Balance encuesta inicial según quienes completan y quienes no completan la investigación

	No completa	Sí Completa.	Dif.	Std. Error.	N
Muestra completa					
Frecuencia de interacciones	17,60	18,19	-0,59	(0,39)	508
Apoyo social percibido	8,43	8,65	-0,22	(0,21)	508
Habilidades de buen trato	15,42	15,56	-0,15	(0,25)	508
Dificultades parentales	4,54	4,78	-0,24	(0,29)	508
Grupo Tratamiento					
Frecuencia de interacciones	17,48	17,50	-0,02	(0,60)	247
Apoyo social percibido	8,34	8,37	-0,03	(0,29)	247
Habilidades de buen trato	15,43	15,28	0,16	(0,34)	247
Dificultades parentales	4,49	4,91	-0,41	(0,42)	247
Grupo Control					
Frecuencia de interacciones	17,73	18,80	-1,07*	(0,51)	261
Apoyo social percibido	8,52	8,89	-0,37	(0,31)	261
Habilidades de buen trato	15,40	15,82	-0,41	(0,37)	261
Dificultades parentales	4,59	4,67	-0,09	(0,40)	261

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Tabla 33: Modelo de sesgo de atrición resultados intermedios (Heckman)

	ITT	Lambda
Frecuencia de interacción	0.267** (0.115)	0.063 (0.339)
Habilidades Parentales - E2P		
Vinculares	0.184* (0.110)	0.168 (0.229)
Protectoras	0.208* (0.107)	0.196 (0.246)
Vinculares	0.184* (0.110)	0.168 (0.229)
Reflexivas	0.234** (0.116)	0.230 (0.252)

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Tabla 34: Modelo de sesgo de atrición resultados primarios (Heckman)

	ITT	Lambda
Desarrollo Infante - ASQ3		
Comunicación	0.066 (0.122)	0.141 (0.283)
Motora fina	0.075 (0.116)	0.214 (0.222)
Motora gruesa	-0.018 (0.126)	0.215 (0.270)
Resolución de problemas	0.017 (0.124)	0.180 (0.234)
Socioindividual	0.140 (0.125)	0.166 (0.278)
Interacción Parental - Piccolo		
Afecto	-0.062 (0.132)	0.233 (0.319)
Responsividad	0.015 (0.132)	0.210 (0.359)
Aliento	0.057 (0.128)	0.400 (0.408)
Enseñanza	0.124 (0.124)	0.776*** (0.231)
Total Piccolo	0.070 (0.127)	0.362 (0.334)

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla 35: Modelo de sesgo de atrición resultados secundarios (Heckman)

	ITT	Lambda
Estrés Parental - PSI4		
Angustia parental	-0.280** (0.110)	-0.245 (0.291)
Interacción disfuncional	-0.225** (0.113)	-0.285 (0.225)
Características difíciles	-0.203* (0.107)	-0.189 (0.260)
Estrés total	-0.285** (0.107)	-0.309 (0.247)
Apoyo social percibido	0.294 (0.292)	0.233 (0.776)

Fuente: elaboración propia. Nota: entre paréntesis los errores estándar robustos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

10 Análisis de costo efectividad

En esta sección se presenta un análisis de costo-efectividad, con el objetivo de estimar la relación entre los costos del programa Crecer Jugando y sus efectos observados. Este análisis permite comparar el costo por unidad de efecto del programa Crecer Jugando con otras alternativas de intervención, aportando evidencia sobre su eficiencia relativa y su potencial escalabilidad en escenarios de decisión. El análisis se desarrolla en los siguientes pasos:

1. Definir los costos de la intervención y determinar el valor presente de los costos considerando una tasa de descuento del 5 % y un horizonte temporal de 5 años.
2. Calcular el costo promedio por día en esos 5 años de intervención.
3. Determinar el ratio de costo efectividad incremental, utilizando para esto el tamaño del efecto ITT.

10.1 Costos de la intervención

El programa Crecer Jugando se divide en dos fases. En la primera, denominada fase de implementación, se concentra una proporción significativa de los costos. Esto se debe a que, durante los primeros semestres, es necesario invertir en la certificación de los equipos, la adquisición de materiales para las intervenciones y un proceso de acompañamiento más intensivo. En la segunda fase, conocida como fase de autonomía, estos costos se reducen considerablemente.

En la Tabla 36 se presenta el desglose de los costos asociados al programa. Se observa que los ítems de mayor peso corresponden a recursos humanos, incluyendo al facilitador/a, monitor/a y coordinador/a, quienes representan el componente más significativo del presupuesto total. En contraste, ciertos costos como la certificación de equipos y la gestión administrativa tienen una mayor incidencia durante el primer año de implementación, reduciéndose en los años siguientes. Una situación similar ocurre con los materiales, cuya adquisición se concentra al inicio del primer año y cuya vida útil se extiende a lo largo de los cinco años del horizonte temporal considerado.

En relación con el costo del espacio físico, la fundación actualmente utiliza salas facilitadas sin costo por las municipalidades donde se implementa el programa. No obstante, con el fin de permitir una comparación más adecuada con otras intervenciones que sí incurren en este tipo de gasto, se realiza una estimación del costo de arriendo correspondiente. Para ello, se considera el valor de arrendar una sala de 30 metros cuadrados en la comuna de El Bosque durante 3 horas diarias, suficiente para la ejecución de los tres ciclos del programa por día (Acurrucados, Exploradores y Creadores). Este costo diario se multiplica por las 10 sesiones que componen cada ciclo y luego por los tres ciclos anuales contemplados en la intervención.

Para los siguientes análisis, se consideran los costos promedio de una comuna cuya capacidad máxima de implementación es de tres ciclos anuales. Dado que en cada ciclo participan tres grupos (Acurrucados, Exploradores y Creadores), y que cada grupo está compuesto, en promedio, por 10 díadas, una comuna puede atender anualmente a 90 familias.

Si bien la capacidad máxima es de 90 familias, el análisis asume, con base en datos de la fundación Infancia Primero, que la cantidad de ciclos anuales aumenta progresivamente. Por lo tanto, durante los primeros tres semestres, se considera que la comuna implementa solo un ciclo semestral. Posteriormente, en los siguientes períodos, se alcanza la capacidad máxima de tres ciclos anuales.

Tabla 36: Listado de costos del programa

Item de costo	Porcentaje
Gestión administrativa	5 %
Certificación de los Equipos	2 %
Kit de Materiales	2 %
Acompañamiento Implementación	25 %
Facilitadora	42 %
Monitora	19 %
Coordinador	4 %
Espacio Físico	2 %

Fuente: elaboración propia en base a información financiera compartida por Fundación Infancia Primero.

La Tabla 37 presenta los costos totales anuales de la intervención, así como el costo promedio anual por día de intervenida. Como se observa, en el año 1 los costos son considerablemente más altos que en los años posteriores, registrándose una caída del 25 % en el año 2 y un 16 % en el año 3, para luego estabilizarse en los años siguientes. A partir de los costos anuales, se calcula el valor presente neto (VPN) de la intervención, aplicando una tasa de descuento del 5 % y una inflación anual del 3 %. Este valor se divide luego por la cantidad total de días atendidos durante el período de cinco años, obteniendo así el costo promedio por día proyectado a cinco años plazo.

Tabla 37: Costos Anuales de la intervención en miles de pesos

	Año 1	Año 2	Año 3	Año 4	Año 5
Costo total	\$ 16.321	\$ 12.179	\$ 10.185	\$ 10.185	\$ 10.185
Días tratadas	60	75	90	90	90
Costo promedio por día	\$ 272,0	\$ 162,4	\$ 113,2	\$ 113,2	\$ 113,2
Costo promedio a 5 años	\$136,9				

Fuente: elaboración propia. Montos en miles de pesos chilenos.

El análisis de costo-efectividad fue realizado considerando el programa en su conjunto, sin diferenciar por modalidad de implementación. Cabe señalar que el cálculo del costo unitario se basa en la modalidad presencial, ya que esta incluye el uso de espacio físico, un componente no presente en la versión online. No obstante, la exclusión de este ítem tiene un efecto marginal sobre el costo total estimado (aproximadamente \$3.000 CLP por día), por lo que no altera sustantivamente los resultados. Dado que la evidencia no muestra efectos diferenciados entre la modalidad presencial y la *online* en la mayoría de los indicadores, y considerando que los costos por participante son similares entre ambas versiones, se considera pertinente utilizar una única estimación de impacto y costo unitario para este ejercicio. En ese sentido, si el programa muestra ser costo-efectivo bajo el escenario más exigente (presencial), su desempeño económico sería incluso más favorable en la modalidad *online*.

10.2 Ratio de costo efectividad

Una vez obtenido el costo unitario, se puede calcular el ratio de costo-efectividad, definido como la relación entre el costo unitario del programa y el tamaño del efecto. Este valor representa el costo promedio necesario para generar un aumento de una desviación estándar en los resultados de una día. Tanto para el tamaño del

efecto como para los ratios de costo efectividad, se presentan los límites inferiores e superiores, tomando en cuenta los límites de los intervalos de confianza al 90 % de los coeficientes ITT calculados.

En la Tabla 38 se presentan los tamaños del efecto estimados para cada una de las dimensiones evaluadas, junto con los ratios de costo-efectividad correspondientes a aquellas dimensiones en las que el estimador ITT resulta estadísticamente significativo. Se destaca que los ratios son inferiores a 600 mil pesos en las habilidades formativas y reflexivas, así como en la angustia parental y el estrés total.

Tabla 38: Ratios de costo efectividad en miles de pesos chilenos

Resultados	Tamaño del efecto			Ratio Costo efectividad		
	Min.		Max.	Max.		Min.
Frecuencia de interaccion	-0,04	0,11	0,26			
E2p - Vinculares	0,03	0,20	0,37	\$ 4.528,8	\$ 692,9	\$ 375,1
E2p - Formativas	0,07	0,24	0,41	\$ 1.931,5	\$ 569,9	\$ 334,3
E2p - Protectoras	0,09	0,20	0,32	\$ 1.504,1	\$ 673,9	\$ 434,2
E2p - Reflexivas	0,03	0,26	0,48	\$ 5.126,3	\$ 536,1	\$ 282,9
Piccolo - Afecto	-0,21	-0,02	0,18			
Piccolo - Aliento	-0,13	0,06	0,25			
Piccolo - Responsividad	-0,21	0,00	0,21			
Piccolo - Enseñanza	0,04	0,21	0,38	\$ 3.593,6	\$ 660,5	\$ 363,7
Piccolo - Total	-0,02	0,09	0,20			
PSI4- Angustia parental	-0,14	-0,26	-0,37	\$ 947,7	\$ 531,9	\$ 369,7
PSI4-Interacción disfuncional	-0,02	-0,20	-0,38	\$ 5.977,9	\$ 676,8	\$ 358,7
PSI4-Características difíciles	-0,06	-0,22	-0,38	\$ 2.209,7	\$ 625,4	\$ 364,3
PSI4- Estrés total	-0,14	-0,27	-0,40	\$ 944,9	\$ 498,5	\$ 338,6
Apoyo social	-0,14	0,02	0,17			

Fuente: elaboración propia. Montos en miles de pesos.

Para poder comparar el ratio de costo efectividad con el de otros programas, en la Tabla 39 se presentan los ratios en dólar estadounidense (USD). En todos los casos, los ratios de costo-efectividad se mantienen por debajo de los USD 750, siendo el valor más bajo el correspondiente a la dimensión de estrés total, con un ratio de USD 533 por unidad de efecto.

Para facilitar la comparación internacional considerando las diferencias de precio existente entre los países, en lugar del tipo de cambio de mercado se puede emplear el factor de conversión por Paridad de Poder Adquisitivo (PPP, por sus siglas en inglés, Purchasing Power Parity)³. Para este cálculo, se utilizan los valores de PPP publicados por el Banco Mundial. Adicionalmente, se incorpora una tasa de inflación del 3 % tanto para Chile como para Estados Unidos, con el fin de homogeneizar los valores comparativos en términos reales.⁴.

³El uso del factor de conversión PPP en lugar del tipo de cambio nominal permite corregir las diferencias en los niveles de precios entre países, proporcionando una medida más precisa del poder adquisitivo relativo.

⁴<https://datos.bancomundial.org/indicador/PA.NUS.PPP>

Tabla 39: Ratios de costo efectividad en dólares a tipo de cambio nominal

Resultados	Tamaño del efecto			Ratio Costo efectividad		
	Min.		Max.	Max.		Min.
Frecuencia de interaccion	-0,04	0,11	0,26			
E2p - Vinculares	0,03	0,20	0,37	\$ 4.838	\$ 740	\$ 401
E2p - Formativas	0,07	0,24	0,41	\$ 2.064	\$ 609	\$ 357
E2p - Protectoras	0,09	0,20	0,32	\$ 1.607	\$ 720	\$ 464
E2p - Reflexivas	0,03	0,26	0,48	\$ 5.477	\$ 573	\$ 302
Piccolo - Afecto	-0,21	-0,02	0,18			
Piccolo - Aliento	-0,13	0,06	0,25			
Piccolo - Responsividad	-0,21	0,00	0,21			
Piccolo - Enseñanza	0,04	0,21	0,38	\$ 3.839	\$ 706	\$ 389
Piccolo - Total	-0,02	0,09	0,20			
PSI4- Angustia parental	-0,14	-0,26	-0,37	\$ 1.012	\$ 568	\$ 395
PSI4-Interacción disfuncional	-0,02	-0,20	-0,38	\$ 6.387	\$ 723	\$ 383
PSI4-Características difíciles	-0,06	-0,22	-0,38	\$ 2.361	\$ 668	\$ 389
PSI4- Estrés total	-0,14	-0,27	-0,40	\$ 1.009	\$ 533	\$ 362
Apoyo social	-0,14	0,02	0,17			

Fuente: elaboración propia. Montos en dólares.

En la Tabla 40 se presenta el factor de cambio PPP, los costos en dólares reales, el total de beneficiarios y el costo promedio por díaada atendida. En base a esa información, el valor presente de los flujos alcanza un monto de USD 133.185. A partir de este cálculo, el costo por díaada atendida se estima en USD 329.

A partir de esta información, la Tabla 41 presenta los ratios de costo-efectividad expresados en dólares considerando las diferencias de precios entre países.

Los ratios calculados nos permiten comparar la efectividad del programa Crecer Jugando, con otros programas con fines similares. Para realizar esta comparación es necesario encontrar programas que hayan sido medidos utilizando los mismos instrumentos que se han utilizado en esta investigación.

Tabla 40: Valor presente en dólares PPA

	Año 1	Año 2	Año 3	Año 4	Año 5
Factor de cambio PPP	436	440	444	449	453
Costo total	37.442	27.668	22.913	22.691	22.471
Total de beneficiarios	60	75	90	90	90
Costo promedio por día	624	369	255	252	250
Valor Presente	133.185				
Costo por día	329				

Fuente: elaboración propia. Monto en dólares.

Tabla 41: Ratios de costo efectividad en dolares PPP

Resultados	Tamaño del efecto			Ratio Costo efectividad		
	Min.		Max.	Max.		Min.
Frecuencia de interaccion	-0,04	0,11	0,26			
E2p - Vinculares	0,03	0,20	0,37	\$ 10.179	\$ 1.557	\$ 843
E2p - Formativas	0,07	0,24	0,41	\$ 4.341	\$ 1.281	\$ 751
E2p - Protectoras	0,09	0,20	0,32	\$ 3.381	\$ 1.515	\$ 976
E2p - Reflexivas	0,03	0,26	0,48	\$ 11.522	\$ 1.205	\$ 636
Piccolo - Afecto	-0,21	-0,02	0,18			
Piccolo - Aliento	-0,13	0,06	0,25			
Piccolo - Responsividad	-0,21	0,00	0,21			
Piccolo - Enseñanza	0,04	0,21	0,38	\$ 8.077	\$ 1.485	\$ 817
Piccolo - Total	-0,02	0,09	0,20			
PSI4- Angustia parental	-0,14	-0,26	-0,37	\$ 2.130	\$ 1.196	\$ 831
PSI4-Interacción disfuncional	-0,02	-0,20	-0,38	\$ 13.436	\$ 1.521	\$ 806
PSI4-Características difíciles	-0,06	-0,22	-0,38	\$ 4.967	\$ 1.406	\$ 819
PSI4- Estrés total	-0,14	-0,27	-0,40	\$ 2.124	\$ 1.120	\$ 761
Apoyo social	-0,14	0,02	0,17			

Fuente: elaboración propia. Montos en dólares.

10.3 Estrés Parental (PSI-4)

Hare & Graziano (2022) realizan un meta análisis para calcular la costo-efectividad del programa Parent-Child Interaction Therapy (PCIT). Este programa consiste en terapia de 16 sesiones aproximadamente ⁵ cuyo objetivo es mejorar la relación entre padres e hijos y reducir comportamientos disruptivos de los infantes. El programan se centra en la enseñanza y en el refuerzo de habilidades de crianza positivas a través de sesiones en vivo.

Los autores realizan un análisis de costo efectividad del programa estandar, y de sus versiones alternativas: PCIT Intensivo (se lleva a cabo en sesiones de 90 minutos, cinco días a la semana durante dos semanas), PCIT en Grupo Pequeño (se realiza con grupos pequeños de familias en sesiones de 60-90 minutos) y PCIT en Grupo Grande (sesiones de 90 minutos una vez a la semana por 8 semanas, con 15-20 familias por grupo).

Para medir el ratio de costo efectividad utilizan, entre otras dimensiones, el estrés total, medido mediante el PSI4. En la Tabla 42 se observa una comparación entre el programa crecer Jugando y el programa PCIT en sus diferentes versiones. Para la comparación se utiliza tanto el costo del programa Crecer Jugando a precios nacionales como a precios internacionales. A partir de esta comparación, se observa que, a precios nacionales, el programa es un 63 % más costo efectivo que la versión intensiva del PCIT, 33 % más que el PCIT Estandar y 4 % más que PCIT en grupo pequeño. Sin embargo, al corregir por PPP, el programa Crecier Jugando es el menos costo efectivo.

Si bien usar el tipo de cambio considerando PPP corrige por las diferencias de precio existentes entre países, no refleja completamente los precios relativos de los bienes y servicios específicos utilizados en un programa. Esto se debe a que la PPP se basa en una canasta estándar de bienes y servicios, mientras que los costos de un programa pueden depender de factores particulares como los salarios o insumos específicos, que varían entre países. Es por esta razón que la literatura recomienda priorizar el tipo de cambio nominal para evitar malas interpretaciones y mantener la transparencia respecto a los costos reales observados (Dhaliwal, Duflo, Glennerster & Tulloch, C. 2020)

Tabla 42: Comparación de costo efectividad en el estrés total

	Costo por familia	Tamaño del efecto	Ratio Costo efectividad
CJ (precios internacionales)	307,78	0,28	1099,2
PCIT Intensivo	756,29	0,89	849,8
PCIT Estandar	842,74	1,21	696,5
PCIT en Grupo Pequeño	530,87	0,98	541,7
CJ (precios nacionales)	146,29	0,28	522,5
PCIT en Grupo Grande	146,50	0,78	187,8

Fuente: elaboración propia. Costo por familia y ratio de costo efectividad está en dólares.

10.4 Habilidades parentales (E2P)

Para realizar un análisis de costo-efectividad del programa Crecer Jugando en el desarrollo de habilidades parentales, se utiliza como referencia el programa Crianza Positiva. Esta iniciativa, diseñada para fortalecer las prácticas de crianza en familias con niños de 0 a 2 años en Uruguay, se implementa en los Centros de Atención a la Infancia y la Familia (CAIF). La intervención consta de un taller de ocho semanas sobre competencias parentales, seguido de un acompañamiento de 24 semanas a través de mensajes de texto y audio.

⁵La cantidad de sesiones es aproximada dado que el programa termina una vez que los cuidadores alcanzan un dominio mínimo de habilidades. Investigaciones indican que las familias tardan entre 13 y 20 semanas en completar el PCIT, por lo que Hare & Graziano utilizan un promedio de ambas duraciones

Bloomfield (2019) evaluó el impacto del componente de mensajería en las familias participantes. Para medir su efecto en las habilidades parentales, utilizó el instrumento E2P, encontrando un impacto estadísticamente significativo de 0,24 DE en las habilidades reflexivas. En cuanto al resto de las dimensiones, la autora realizó ajustes en las escalas, por lo que no se incluirán en el análisis.

Para calcular los costos de esta intervención se utiliza como referencia los costos del programa Crecer Jugando, dado que en ambas intervenciones se realiza el envío de mensajes a las familias vía Whatsapp.

Para calcular los costos de la componente mensajería del programa Crecer Jugando, se observan las horas dedicadas por el equipo a cada una de las labores que realiza el programa. En base a esto se identifica que de las 16 semanas dedicadas al programa, 9 de ellas son exclusiva para la intervención presencial. De esta forma se determina que un 44 % de los costos se pueden asignar al componente de mensajería.

Teniendo esta estimación, se elimina del listado de costos el costo asociado al arriendo de espacios físicos y de los materiales necesarios para realizar la intervención presencial. Para todo el resto de los costos (certificación, acompañamiento, gestión administrativa) se asume que un 44 % de estos se puede asignar al componente mensajería.

A partir de este supuesto se obtiene que el componente mensajería tiene costos anuales como los presentados en la Tabla 43. Luego, con esta información se calcula el costo por cada intervención para 10 díadas.

A continuación, considerando que el programa Crecer Jugando tiene 10 semanas de mensajes y que por su parte el programa Crianza Positiva 24 semanas, se multiplica el costo anual del primero por 24/10, obteniendo así el costo anual de un ciclo de intervención. Adicionalmente, asumiendo que el programa Crianza Positiva puede realizar máximo dos intervenciones anuales, se obtiene un costo anual y el costo unitario, que es el presentado en la Tabla 43.

Tabla 43: Costos componente mensajería

	Año 1	Año 2	Año 3	Año 4	Año 5
Componente mensajería CJ	\$ 6.477	\$ 5.207	\$ 4.335	\$ 4.335	\$ 4.335
Componente mensajería CP	\$ 5.182	\$ 3.332	\$ 2.312	\$ 2.312	\$ 2.312
Valor Presente de los Costos CP	\$ 14.623				
Díadas totales atendidas	100				
Costo unitario CP	\$ 146,2				

Fuente: elaboración propia. Monto en pesos chilenos.

Con esta información, la Tabla 44 muestra que el programa Crecer Jugando es un 12 % más costo-efectivo que Crianza Positiva. Esta diferencia se explica principalmente porque Crianza Positiva tiene un mayor tiempo de intervención, lo que incrementa sus costos, y les permite atender a una menor cantidad de familias por año.

Es importante señalar que, si bien este resultado aporta indicios sobre la costo-efectividad del programa, el análisis se basa en ciertos supuestos. El más relevante es que se asume un costo por sesión de mensajería igual para ambos programas, lo que podría no ser preciso si existieran diferencias en la estructura organizacional, las tarifas horarias, la productividad de los facilitadores, entre otras variables.

Adicionalmente, para garantizar la comparabilidad entre ambos programas, el cálculo se realizó bajo la premisa de que cada intervención atiende a 10 familias. No obstante, una intervención basada exclusivamente en el envío de mensajes tiene el potencial de alcanzar un mayor número de familias por ciclo, lo que podría influir en su costo-efectividad de ambos programas.

Tabla 44: Comparación de costo efectividad en habilidades reflexivas

	Costo por familia	Tamaño del efecto	Ratio Costo efectividad
CJ (pesos chilenos)	\$ 136.930	0,255	\$ 536.982
CP (pesos chilenos)	\$ 146.238	0,243	\$ 601.801

Fuente: elaboración propia. Costo por familia y ratio de costo efectividad está en pesos chilenos.

11 Conclusiones y desafíos

Aportes generales de la evaluación

La evaluación de impacto del programa Crecer Jugando entrega evidencia sobre los efectos del programa en sus resultados esperados según su modelo de cambio. Como intervención de carácter preventivo y promocional, Crecer Jugando se orienta a fortalecer las habilidades parentales y promover vínculos significativos entre cuidadores y niños/as, con el objetivo de favorecer el desarrollo infantil en sus primeras etapas. A partir de esta lógica, el programa también busca reducir factores de riesgo como el estrés parental, considerado un resultado secundario relevante por su influencia en la calidad de la interacción cuidador–niño/a.

Esta evaluación, realizada mediante un diseño experimental, permite no solo identificar resultados atribuibles al programa, sino también avanzar en una comprensión más profunda de su efectividad bajo distintas condiciones. En particular, entrega evidencia sobre cómo varía su impacto según el rango etario de los niños/as participantes, la modalidad de implementación y el nivel de participación efectiva. La incorporación de análisis de efectos heterogéneos y marginales permite identificar umbrales de intensidad a partir de los cuales se observa mayor impacto, aportando información clave para la toma de decisiones sobre el diseño operativo del programa, como el número mínimo de sesiones necesarias. Estos hallazgos fortalecen la base para orientar ajustes estratégicos y proyectar su escalabilidad con criterios de costo-efectividad.

Asimismo, la evaluación ofrece una oportunidad para reflexionar sobre aquellos ámbitos en los que no se identificaron efectos significativos, lo que puede deberse tanto a aspectos del diseño de la intervención como a limitaciones en las estrategias de medición o en la sensibilidad de los instrumentos utilizados.

Resultados generales de la evaluación de impacto

En términos generales, los análisis muestran que el programa genera efectos positivos en las habilidades parentales evaluadas con el instrumento E2P, con mayor intensidad en las dimensiones reflexiva y formativa. También se identifican reducciones significativas en el estrés parental, donde consistentemente, tanto en el estrés total como en todas sus dimensiones asociadas, se observan reducciones que llegan a estar por sobre las 0.4 desviaciones estándar.

A esto se suma un efecto observado en el desarrollo infantil, específicamente en el dominio socio individual. Si bien este hallazgo es consistente con la hipótesis causal del programa -que plantea que la mejora en el vínculo y las prácticas de crianza debiese favorecer el desarrollo infantil-, en general no se identificaron efectos significativos en otras dimensiones del desarrollo evaluadas. Este resultado es esperable si se considera que el desarrollo infantil es un proceso acumulativo y de más largo plazo dentro de la teoría de cambio del programa, y que la medición fue realizada inmediatamente después del cierre de la intervención, sin dar espacio suficiente para que posibles efectos se manifesten de forma más clara. En este sentido, se recomienda considerar instancias de seguimiento longitudinal que permitan evaluar con mayor precisión los impactos del programa en esta área.

En cuanto al fortalecimiento del vínculo cuidador/a–niño/a, los resultados de impacto muestran un efecto positivo en la dimensión de enseñanza del instrumento PICCOLO. Sin embargo, los análisis de sensibilidad aplicados a este resultado, evidencian la presencia de sesgo de selección, lo que sugiere que ciertas características de quienes aceptaron ser grabados podrían estar influyendo en el efecto observado. Este hallazgo contrasta con lo encontrado en los otros instrumentos de medición utilizados en la evaluación, como el E2P o el cuestionario de estrés parental, donde no se identifican indicios de sesgo.

A lo anterior se suma que el PICCOLO, diseñado originalmente con fines formativos más que evaluativos, puede presentar limitaciones para detectar cambios atribuibles a intervenciones de corta duración. Por estas razones,

los efectos observados en esta dimensión deben interpretarse con cautela.

De todas formas, esto no permite concluir que el programa no tiene efectos en el vínculo. El E2P -instrumento que incluye dimensiones asociadas a la calidad de la relación cuidador-niño/a, como sensibilidad, calidez emocional e involucramiento- muestra efectos positivos, y además existe una correspondencia general entre los cambios observados en habilidades parentales y reducción del estrés, dos dimensiones que tienden a estar relacionadas con la calidad del vínculo. Esto refuerza la necesidad de seguir mejorando las estrategias para evaluar este tipo de dimensiones en futuras investigaciones

Interpretación de los efectos: ITT y LATE como estimaciones complementarias

Como se evidencia en los resultados de fidelidad, se observa una tasa considerable de abandono entre las díadas que inician la intervención, ya que muchas de ellas no continúan una vez comenzado el programa. Según los estándares definidos, cerca del 60 % de las díadas inscritas alcanzan la “dosis mínima” de participación. Este patrón tiene implicancias importantes para la interpretación de los resultados de impacto.

Las estimaciones LATE (Local Average Treatment Effect) permite aproximar el efecto del programa sobre quienes efectivamente participan de la intervención, entregando información especialmente útil para la toma de decisiones sobre su diseño, implementación y ajustes operativos. Este enfoque permite identificar el efecto real en las condiciones en que el programa se experimenta tal como fue planificado, y es particularmente relevante para comprender qué tan efectivos son sus contenidos y dinámicas cuando se logra una participación sostenida.

Por su parte, el enfoque intention-to-treat (ITT), que incluye tanto a quienes participan como a quienes no completan la intervención, entrega una estimación del impacto del programa en condiciones reales de implementación, considerando todo el esfuerzo institucional y operativo necesario para ponerlo en marcha. Es por esto que el estimador ITT sigue siendo fundamental cuando se busca evaluar el desempeño del programa como política pública y es el estimador adecuado para análisis de costo-efectividad. En conjunto, ambas estimaciones-ITT y LATE- ofrecen miradas complementarias que permiten comprender no solo el potencial de la intervención, sino también su comportamiento al ser escalada o replicada en distintos contextos.

Efectos según intensidad de participación

Uno de los aportes metodológicos más relevantes de esta evaluación es la estimación del efecto marginal según la cantidad de sesiones asistidas por las díadas. Los resultados muestran que no se observan efectos significativos en las primeras sesiones, pero a partir de la tercera sesión comienzan a evidenciarse impactos positivos tanto en las habilidades parentales (especialmente en las dimensiones reflexiva, formativa y vincular) como en la reducción del estrés parental.

Estos efectos se consolidan entre la cuarta y quinta sesión, alcanzando su mayor magnitud en ese rango, y tienden a estabilizarse a partir de la sexta sesión. Esto sugiere que existe un umbral mínimo -alrededor de 3 o 4 sesiones- a partir del cual se genera impacto, y que la cantidad óptima para maximizar beneficios estaría entre 3 y 6 sesiones. Esta información puede ser especialmente útil para orientar estrategias de acompañamiento o seguimiento focalizado, priorizando a aquellas díadas que se encuentran cerca de alcanzar el umbral mínimo de participación. En estos casos, una intervención oportuna podría marcar la diferencia en términos de impacto, más que en participantes que aún no han iniciado el proceso o que ya cuentan con una asistencia mínima lograda o sostenida.

Efectos diferenciados por grupos y modalidades

Los efectos del programa varían según el grupo etario al que pertenecen las díadas, en línea con los énfasis diferenciados que propone la intervención para cada etapa del desarrollo infantil. En el grupo Exploradores (13

a 24 meses), se identifican mayores efectos en habilidades parentales en la dimensión reflexiva y protectora, así como en dimensiones asociadas al vínculo cuidador-niño/a, como responsividad y enseñanza. Además, este grupo presenta mejoras significativas en la dimensión de comunicación del desarrollo infantil, en comparación con el grupo Acurrucados. Estos resultados se mantienen al controlar por características de las díadas, lo que sugiere efectos diferenciados atribuibles a la intervención en esa etapa.

También se observan efectos mayores en la reducción del estrés parental en los grupos Exploradores y Creadores respecto a Acurrucados, diferencia que persiste tras ajustar por variables observables. Este patrón podría estar asociado a mayores niveles de sobrecarga o complejidad en las etapas más avanzadas de la crianza, donde el programa puede resultar especialmente útil.

En el resto de las dimensiones evaluadas, no se identifican diferencias estadísticamente significativas entre los grupos etarios. Esto sugiere que, si bien hay efectos diferenciados en algunas variables, el programa no muestra un patrón sistemático de mayor efectividad en un tramo etario específico, y tiende a operar de forma relativamente estable en distintas etapas del desarrollo temprano. La concentración de ciertos efectos los exploradores puede abrir una línea de análisis futura sobre una posible mayor costo-efectividad del programa en esas edades. No obstante, esta hipótesis requiere ser explorada en mayor profundidad, idealmente mediante un diseño específico que permita evaluarla con ese propósito.

En relación con la modalidad de implementación, se observa que tanto en formato presencial como en formato online se registran efectos positivos en los resultados evaluados. En la mayoría de los casos, no se identifican diferencias significativas entre modalidades, lo que respalda la lógica de implementación flexible que define el modelo actual del programa. La única excepción corresponde al apoyo social percibido, donde se observa un efecto mayor en la modalidad presencial, posiblemente asociado a las oportunidades de interacción directa entre cuidadoras. Sin embargo, estos resultados deben interpretarse con cautela, dado que el grupo que participó en modalidad online es considerablemente más pequeño (33 díadas), lo que limita la potencia estadística de los análisis y, por tanto, la posibilidad de identificar diferencias en otras dimensiones.

En conjunto, estos hallazgos apoyan la viabilidad técnica del modelo flexible de implementación, y sugieren que el programa mantiene su efectividad en distintos contextos y formatos. Esta flexibilidad representa una fortaleza al momento de pensar en su expansión, ya que permite adaptarse a las condiciones locales sin comprometer los resultados esperados.

Costo-efectividad

Desde el análisis de costo-efectividad, los resultados posicionan al programa Crecer Jugando como una intervención eficiente en relación con el nivel de inversión necesario para generar impactos en dimensiones prioritarias.

El costo promedio por díada beneficiaria se estima en aproximadamente \$137.000 CLP, y los ratios de costo-efectividad se sitúan en torno a \$600.000 CLP por desviación estándar, dependiendo de la dimensión considerada. En particular, se destacan los ratios de \$498.000 CLP para el efecto en estrés parental total y de \$536.000 CLP para habilidades reflexivas.

Si bien la disponibilidad de referentes comparables en algunas dimensiones es limitada, estas cifras muestran un desempeño competitivo frente a intervenciones internacionales de referencia, especialmente cuando se consideran precios locales.

Esto refuerza la viabilidad económica del programa y aporta evidencia favorable respecto a su potencial escalabilidad. Además, el hecho de que no se observen diferencias significativas en la efectividad entre las modalidades presencial y *online*, junto con costos unitarios similares - e incluso levemente más bajos en la versión remota-,

destaca la adaptabilidad del modelo sin comprometer su eficiencia. Esta combinación de efectividad técnica, flexibilidad operativa y uso razonable de recursos constituye una base propicia para pensar en su sostenibilidad y expansión a mayor escala.

A ello se suma un componente central del diseño del programa: la instalación de capacidades en los equipos locales y en las instituciones territoriales. Este enfoque no solo favorece la apropiación y continuidad de la intervención una vez finalizado el acompañamiento externo, sino que refuerza el carácter preventivo y comunitario del modelo. En este marco, resulta especialmente relevante contar con evidencia de impacto en un programa que apuesta por el fortalecimiento de capacidades locales como vía para lograr efectos sostenibles en el tiempo.

Consideraciones sobre la robustez de los resultados

Los análisis realizados para evaluar el posible sesgo asociado a la pérdida de casos a lo largo del estudio muestran que, en general, la atrición observada no compromete la validez de los resultados estimados. El uso del modelo de selección de Heckman permitió contrastar si las características de las díadas que permanecen en el estudio difieren sistemáticamente de aquellas que no completaron las mediciones.

En la mayoría de los instrumentos aplicados, no se identificó evidencia de sesgo de selección atribuible a la atrición, lo que respalda la validez de los efectos reportados. La única excepción corresponde al instrumento PICCOLO, donde sí se detecta sesgo de selección asociado a la participación en la grabación de interacciones. Esto implica que los efectos estimados con ese instrumento deben interpretarse con cautela, ya que pueden estar influenciados por características no observadas de quienes aceptaron ser grabados.

En conjunto, estos resultados permiten afirmar que los efectos reportados para el conjunto principal de variables no se ven afectados por sesgos asociados a la pérdida de casos, y pueden considerarse representativos de la población evaluada. Además, en ciertos análisis se observan patrones que sugieren que las díadas que completan la intervención tienden a presentar condiciones más desafiantes que aquellas que se retiran, lo que podría implicar que las estimaciones obtenidas —al incluir principalmente a quienes enfrentan mayores niveles de estrés o riesgo— reflejan efectos conservadores. En ese sentido, es posible que el impacto real del programa sea igual o incluso levemente superior al observado en esta evaluación.

Proyección y desafíos futuros

Los resultados de esta evaluación aportan evidencia sobre los efectos inmediatos del programa Crecer Jugando en los cuidadores y niños/as que participan, pero también abren interrogantes importantes sobre su impacto en el mediano y largo plazo. En particular, el desarrollo infantil es un proceso progresivo, cuyos efectos acumulativos pueden expresarse con mayor claridad una vez que los niños y niñas transitan hacia nuevas etapas de aprendizaje y socialización. En este sentido, un desafío pertinente sería avanzar hacia un seguimiento longitudinal, que permitan observar la trayectoria de los participantes más allá del cierre inmediato de la intervención. Este tipo de información no solo permitiría evaluar la sostenibilidad de los cambios observados, sino también afinar la propuesta metodológica del programa para maximizar sus efectos a lo largo del ciclo vital.

12 Referencias Bibliográficas

- Aracena, M., Gómez, E., Undurraga, C., & Leiva, L. (2013). Validación del Índice de Estrés Parental Versión Abreviada (PSI-SF) en población chilena
- Bloomfield, J., Balsa, A., & Cid, A. (2023). Using behavioral insights in early childhood interventions: the effects of Crianza Positiva e-messaging program on parental investment. *Review of Economics of the Household*, 21(1), 95-130.
- Centro CUIDA. (2022, agosto 8). Centro CUIDA dicta curso “Evaluando la parentalidad a través de la escala PICCOLO”. Pontificia Universidad Católica de Chile. <https://cuida.uc.cl/centro-cuida-dicta-curso-evaluando-la-parentalidad-a-traves-de-la-escala-piccolo/>
- Clarke, D., Romano, J., & Wolf, M. (2019). The Romano-Wolf Multiple Hypothesis Correction in Stata. IZA Institute of Labor Economics.
- Cohen, E., & Franco, R. (1990). Evaluación de proyectos sociales: la utilización del análisis costo-efectividad.
- Dhaliwal, I., Duflo, E., Glennerster, R., & Tulloch, C. (2020). Cost-effectiveness analysis. Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*
- Fox, S. E., Levitt, P., & Nelson, C. A. (2010). How the timing and quality of early experiences influence the development of brain architecture. *Child Development*, 81(1), 28-40.
- Gertler, P. J., Martínez, S., Premand, P., Rawlings, L. B., & Veermersch, C. M. J. (2016). La evaluación de impacto en la práctica (2ª ed.). Grupo Banco Mundial / Banco Interamericano de Desarrollo.
- Gómez Muzzio, E., & Contreras Yévenes, L. (2019). Manual Escala de Parentalidad Positiva – E2P v.2. Fundación América por la Infancia.
- Hawk, B. N., Timmer, S. G., Armendariz, L., Boys, D. K., Urquiza, A. J., & Garcia, E. F. y. (2022). Improving children’s behavior in seven sessions: a randomized controlled trial of parent-child care (pc-care) for children aged 2-10 years. *Child Psychiatry & Human Development*, 55(2), 336-349.
- Heckman, J. (1976) The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement*, 5, 475-492.
- Hare, M. M., & Graziano, P. A. (2021). The cost-effectiveness of parent-child interaction therapy: Examining standard, intensive, and group adaptations. *Administration and Policy in Mental Health and Mental Health Services Research*, 48(3), 499-513.
- Iválua (2009) Guía práctica 4: Evaluación de la implementación. Instituto Catalán de Evaluación de Políticas Públicas
- Jeong, J., Franchett, E., Oliveira, C. V. R. d., Rehmani, K., & Yousafzai, A. K. (2021). Parenting interventions to promote early child development in the first three years of life: a global systematic review and meta-analysis. *PLOS Medicine*, 18(5), e1003602.

- Prieto, L., Sacristán, J. A., Antoñanzas, F., Rubio-Terrés, C., Pinto, J. L., & Rovira, J. (2004). Análisis coste-efectividad en la evaluación económica de intervenciones sanitarias. *Medicina clínica*, 122(13), 505-510.
- Maulik, PK. y Darmstadt, GL. (2009). "Community-based interventions to optimize early childhood development in low resource settings" en *Journal of Perinatology*, 29, 531-542.
- NICHD Early Child Care Research Network. Child-care effect sizes for the NICHD Study of Early Child Care and Youth Development. *Am Psychol*. 2006 Feb-Mar;61(2):99-116.
- Nores, M. y Barnett, S. (2010). "Benefits of early childhood interventions across the world: (Under) Investing in the very young" en *Economics of Education Review*, 29, 271-282.
- Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V. K., & Christiansen, K. (2013). PICCOLO: Evaluación de Interacciones Parentales con Niños: Lista de Observaciones Vinculadas a Resultados.
- Squires, J., Twombly, E., Bricker, D., & Potter, L. (2009). ASQ-3: Ages & Stages Questionnaires, Third Edition. Paul H. Brookes Publishing Co.
- Staiger, D., Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557-586.

