

Evaluación de Impacto del programa Crecer Jugando

Informe metodológico

10 de septiembre 2024

Contenido

I.	Introducción	4
II.	Resumen del proyecto.....	4
III.	Protocolo metodológico	6
1.	Antecedentes y justificación	6
2.	Intervención del Programa Crecer Jugando.....	7
3.	Modelo de cambio.....	9
4.	Diseño de la evaluación y plan de análisis	12
4.1	Intención de tratamiento (ITT).....	12
4.2	Plan de análisis ITT	13
4.3	Efecto Promedio de Tratamiento Local (LATE)	15
4.4	Plan de análisis LATE	15
4.5	Análisis adicionales	16
4.5.1	Efectos de interacción	16
4.5.2	Análisis de robustez	17
5.	Cálculos Efecto Mínimo Detectable.....	22
6.	Universo y participantes	23
7.	Análisis de costo efectividad.....	24
8.	Análisis de fidelidad	25
9.	Consideraciones éticas	26
10.	Equipo.....	29
IV.	Informe estadístico y estado de avance de la evaluación.....	30
1.	Aleatorización y terreno	30
2.	Actualización y proyección del Efecto Mínimo Detectable.....	35
3.	Caracterización de los participantes (preliminar a la fecha)	36
5.1	Análisis de la totalidad de participantes que inician el proceso de evaluación (preliminar a la fecha)	36
5.1.1	Balance de grupos	36
5.1.2	Cumplimiento (totalidad de participantes).....	38
5.2	Análisis grupo de participantes que termina el proceso de evaluación (preliminar a la fecha)	40
5.2.1	Balance de grupos	40
5.2.2	Cumplimiento (grupo que termina el proceso de evaluación)	42

4.	Conclusiones y desafíos del estado de avance	44
V.	Bibliografía.....	45

I. Introducción

El presente informe corresponde al informe metodológico preliminar comprometido en el marco del convenio tripartito IMPACTO EDU-LAC para el desarrollo de Evaluaciones de Impacto en América Latina y el Caribe de SUMMA, entre esta última institución, Infancia Primero y Focus, para la evaluación de impacto del Programa Crecer Jugando de Fundación Infancia Primero.

Este informe contiene el protocolo metodológico junto con los requerimientos del plan de análisis estadístico (SAP)

II. Resumen del proyecto

Cuadro 1. Resumen del proyecto

TÍTULO DEL PROYECTO	Evaluación de Impacto Programa Crecer Jugando
DESARROLLADOR (INSTITUCIÓN)	Fundación Infancia Primero
EVALUADOR (INSTITUCIÓN)	Estudios y Consultorías Focus
INVESTIGADOR(ES) PRINCIPAL(ES)	Isidora Fuenzalida
DISEÑO EXPERIMENTAL	Experimento aleatorio de dos brazos con asignación aleatoria a nivel de diadas (cuidador/a y niño/a)
TIPO DE EXPERIMENTO	Efectividad
RANGO DE EDAD DE LOS NIÑOS/AS	0 a 48 meses
NÚMERO DE COMUNAS Y GRUPOS	9 comunas y 21 grupos
NÚMERO DE DIADAS	460 es el número final a alcanzar. Para este informe se consideran las 296 diadas aleatorizadas
MEDIDA DE OUTCOMES PRIMARIOS Y FUENTES	● Respuesta sensible (PICCOLO) ● Desarrollo de habilidades lingüísticas, motoras, de resolución de problemas y sociales (ASQ-3)

RESULTADOS INTERMEDIOS Y FUENTE	● Frecuencia de interacciones con niño/a (Encuesta Infancia Primero) ● Habilidades de buen trato (Encuesta Infancia Primero) ● Habilidades parentales (E2P)
MEDIDAS DE OUTCOMES SECUNDARIOS Y FUENTES	● Apoyo social cuidadores (Encuesta Infancia Primero) ● Estrés parental (PSI-4-SF)

Fuente: elaboración propia

III. Protocolo metodológico

1. Antecedentes y justificación

La evidencia científica hace referencia a la importancia de la primera infancia para el desarrollo cerebral y la adquisición de herramientas y habilidades claves para el futuro. 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, diversos estudios señalan que la crianza, en el marco del cuidado infantil, es reconocida como el primer proceso de socialización de los niños y niñas en el curso de su desarrollo, siendo ésta un elemento central para que todo niño y niña pueda alcanzar su máximo potencial (Uniceg.org)

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 acumulada por décadas en el mundo entero, 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.

Con la evidencia acumulada sobre la correcta implementación en procesos de formación y transferencia del modelo, resultados pre y post en participantes que se mantienen en el programa, además de contar con resultados significativos de evaluaciones anteriores, realizar una evaluación de impacto es un paso natural para que esta intervención siga su curso hacia el escalamiento. Análisis externos han visto el potencial de escalamiento de Crecer Jugando y la evaluación de impacto es una respuesta para enfrentar ese desafío.

Para alcanzar la meta de incluir este programa en espacios de incidencia pública, logrando una mayor eficiencia en la cobertura, se requiere contar con un respaldo sobre el buen uso de los recursos. En esta línea, la evaluación de impacto también pasa a ser un insumo de vital importancia.

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 incluso con evidencia estadísticamente significativa para algunos 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.

A la fecha, se han realizado 5 evaluaciones del Programa Crecer Jugando a partir del trabajo con equipos externos y expertos/as en la materia, que han permitido analizar los diversos procesos de cambio que ha experimentado. Toda esta información se encuentra disponible en la página web de la Fundación.

Con el objetivo de evaluar el impacto de este Programa, se solicita a Estudios y Consultorías Focus realizar estudio que permita identificar si su implementación genera algún impacto en los resultados intermedios, finales y secundarios del diseño del programa, 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. Intervención del Programa Crecer Jugando

Crece Jugando es un programa de apoyo a la parentalidad y al desarrollo infantil temprano que tiene el objetivo de fortalecer el vínculo y la interacción entre el niño o la niña y su cuidador principal (diadas) como medio para potenciar el desarrollo infantil, al entregar competencias a los cuidadores y potenciar las redes de apoyo entre ellos. El público objetivo del programa son niños y niñas 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 diadas. 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, Crece 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, tras el fin de las restricciones sanitarias, 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 supervisa y acompaña de manera constante las actividades desarrolladas por ambos actores.

3. Modelo de cambio

La Teoría de Cambio del programa describe de manera causal los resultados esperados (outcomes 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.

- **Resultados Intermedios**

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.
- **Aumento del conocimiento y habilidades de buen trato:** Los cuidadores adquieren conocimientos y habilidades para vincularse positivamente con sus hijos e hijas.
- **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.

Estos resultados intermedios son cruciales, ya 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 & Barnett, 2009; Maulik & Darmstadt, 2009).

- **Outcomes primarios**

Al generar cambios en estas variables intermedias mencionadas, se espera **mejorar el vínculo entre el/la cuidador/a y el niño o niña**. Esto, a su vez, debería llevar a **mejoras en el desarrollo del niño o niña en áreas como la psicomotricidad, el lenguaje, las habilidades cognitivas y socioemocionales**, identificadas como los outcomes 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).

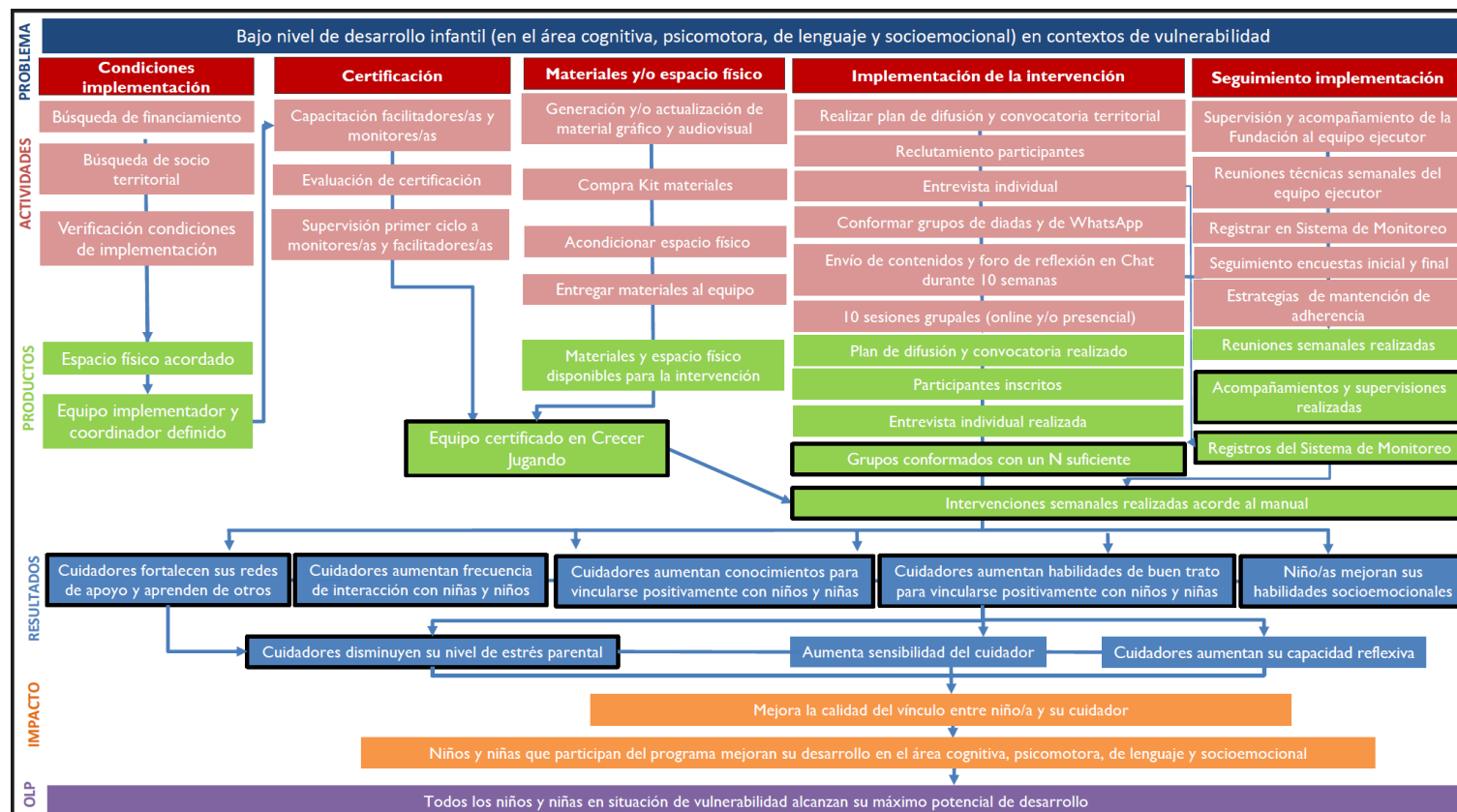
- **Outcomes Secundarios**

Además de los outcomes primarios, el diseño del programa permite identificar outcomes 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.

Figura 1: Teoría de cambio Crecer Jugando



4. Diseño de la evaluación y plan de análisis

4.1 Intención de tratamiento (ITT)

Para evaluar el programa Crecer Jugando se propone una evaluación experimental. En este tipo de investigación se asigna de manera aleatoria al grupo que recibirá el tratamiento (grupo tratamiento) y al grupo que no lo recibirá (grupo control). De esta forma nos aseguramos que la única diferencia entre ambos grupos es que uno de ellos recibe el tratamiento y el otro no. Por lo tanto, las diferencias observables posterior a la aplicación del tratamiento, son únicamente explicados por el programa, asegurando la validez interna de la investigación.¹

Para este caso la unidad de aleatorización son las diadas, conformada por un/a niño/a y un/a cuidador/a. Estas deben cumplir con las siguientes condiciones:

1. Haberse inscrito voluntariamente para participar del programa.
2. No haber participado anteriormente del programa
3. No estar recibiendo ninguna otra intervención relacionada con habilidades parentales.
4. No tener antecedentes de diagnóstico asociado al desarrollo psicomotor que puedan afectar su participación en el programa.
5. Haberse inscrito antes de que se complete el número máximo de diadas atendibles. Aquellas diadas que se inscriban después de que se alcance este máximo no serán elegibles para la intervención en ese ciclo. Sin embargo, podrán ser consideradas para la aleatorización en ciclos posteriores, si continúan interesados en postular.

En relación a la muestra, esta está conformada por las diadas que se inscribieron de manera voluntaria en el programa. Por consiguiente, se trata de un muestreo no aleatorio, lo que implica que no podemos garantizar su validez externa.²

La ecuación para medir el impacto del programa será la siguiente:

$$Y_i = \beta_0 + \beta_1 CJ_i + \beta_n X_{ni} + \mu_i \quad (1)$$

Donde Y_i es un *outcome* de interés, CJ_i es una variable dummy que toma el valor 1 si es que la diada i fue tratada por el programa Crecer Jugando y β_1 el efecto del programa. Por su parte, X_{ni} representa una matriz de n características de la diada, tales como:

- Nivel educacional cuidador/a
- Género cuidador/a
- Género niño/a
- Edad del cuidador/a

1 Asegura que el efecto observado en la variable de resultado se debe a algún aspecto del Tratamiento y no a otros factores.

2 Asegura que el resultado de la muestra va a representar el resultado de la población, dado un nivel de error de muestreo.

- Edad del niño/a
- Nacionalidad
- Relación diada (padre, madre, tío/a, etc.)
- ¿Niño/a asiste a establecimiento educacional?
- ¿Niño/a tiene alguna discapacidad?
- Situación laboral cuidador/a
- Número del hijo/a (1ro, 2do, 3ro, etc.)

Por último, μ_i es un componente de error.

4.2 Plan de análisis ITT

El análisis de resultados se llevará a cabo en 2 grandes procesos. En una primera instancia se realizará una estimación de resultados del programa sobre dos *outcomes* primarios, dos secundarios y tres intermedios. Para permitir la comparabilidad con otros estudios de similar naturaleza, todos los resultados serán expresados en términos del tamaño del efecto en desviaciones estándar. Adicionalmente todos serán evaluados a un 90%, 95% y 99% de confianza.

Si bien, se puede observar que la investigación cuenta con múltiples *outcomes*, podría ser pertinente realizar una corrección de Bonferroni³. Sin embargo, como ya se mencionó en la sección 3, en la literatura se observa una alta correlación entre los *outcomes* que forman parte de este estudio, por lo que no es recomendable llevar a cabo correcciones de este tipo para esta investigación. Por lo tanto, para este caso se realizarán pruebas de múltiples hipótesis usando la metodología Romano-Wolf, que toma en cuenta la correlación que pudiese existir entre los distintos instrumentos que miden resultados. (Clarke et al. 2019).

4.2.1 Análisis Outcomes Primarios

En cuanto a la medición de la respuesta sensible, el PICCOLO se estructura en cuatro dimensiones: (1) afecto, (2) responsividad, (3) aliento o ánimo y (4) enseñanza. Cada una de estas dimensiones se evalúa mediante ítems que se califican de 0 a 2, otorgando un puntaje máximo de 14 puntos para las tres primeras dimensiones y de 16 puntos para la dimensión de enseñanza. Para medir el efecto del programa se realizará un análisis diferenciado por dimensión, tomando los puntajes de cada una de ellas como una variable continua.

El desarrollo de habilidades lingüísticas se mide utilizando el instrumento ASQ-3, que se divide en cinco dominios: (1) comunicación, (2) motricidad gruesa, (3) motricidad fina, (4) resolución de

³ La corrección de Bonferroni consiste en ajusta el nivel de significancia dividiéndolo por el número de pruebas a realizar, disminuyendo con esto la probabilidad de cometer error tipo 1(rechazar hipótesis nula cuando era verdadera)

problemas y (5) habilidades personales y sociales. En este instrumento, los cuidadores deben responder a cada ítem con “sí” (puntuación = 10), “a veces” (puntuación = 5) o “todavía no” (puntuación = 0)⁴. Cada dominio contiene 30 ítems, por lo que la puntuación total varía de 0 a 60 por dominio. Para el análisis de resultados se realizará un análisis diferenciado por dimensión, tomando el puntaje de cada dimensión como variable dependiente

En ambos casos se evaluará inicialmente el efecto ITT (Intención de Tratamiento) sobre los *outcomes* primarios mediante regresiones lineales donde las variables dependientes serán los puntajes del PICCOLO y los puntajes estandarizados del ASQ-3. En primera instancia, se utilizará la participación en el programa como única variable independiente. Posteriormente, como un primer análisis de robustez, se agregarán progresivamente covariables para identificar si el efecto varía a medida que se controla por diferentes características observables tanto de los/as cuidadores/as y niños o niñas.

4.2.2 Análisis Outcomes secundarios

En relación con los resultados secundarios, el PSI4-SF se divide en tres dimensiones: (1) Estrés Parental, (2) Interacción Disfuncional Padre-Hijo y (3) Características Difíciles del Niño. Cada una de estas dimensiones está compuesta por 12 ítems, con un puntaje máximo de 60. Para el análisis de resultados, se realizará un análisis diferenciado por dimensión, considerando el puntaje como una variable continua.

4.2.3 Resultados intermedios

En cuanto a las habilidades parentales, el E2P se divide en cuatro competencias parentales: Vinculares, Formativas, Protectoras y Reflexivas. El instrumento consta de 60 ítems para todos los rangos de edad, cada uno con 5 opciones de frecuencia, donde “Nunca” asigna cero puntos y “Siempre” asigna 4 puntos. La única excepción es para los niños menores de 3 meses, donde “Nunca”, “Casi Nunca” o “A veces” asignan 2 puntos, “Casi Siempre” asigna 3 puntos y “Siempre” asigna 4 puntos. Para el análisis, se realizará un análisis diferenciado por competencia parental, considerando el puntaje como una variable continua.

En relación con las habilidades de buen trato y la frecuencia de interacción, se utilizará un método de diferencias en diferencias, aprovechando la línea base disponible, para identificar si el programa impacta estos ámbitos. Cabe destacar que tanto las habilidades de buen trato como la frecuencia de interacción se miden en la encuesta inicial, y están compuestas por una serie de preguntas que apuntan a identificar la frecuencia de prácticas realizadas por las díadas. Por lo tanto, en este caso, el análisis de resultados se realizará por pregunta, evaluando el nivel de impacto que tiene el programa en las diferentes prácticas llevadas a cabo por las díadas evaluadas.

⁴ Cuando faltan uno o dos ítems por dominio, deben ser reemplazados por la puntuación promedio del niño o niña en ese dominio

4.3 Efecto Promedio de Tratamiento Local (LATE)

Dado que la participación efectiva en el tratamiento para las diadas del grupo tratamiento es voluntaria (no se puede obligar a nadie a participar si es que es seleccionada), los efectos encontrados en la fórmula (1) corresponde a lo que se conoce como el efecto de la Intención de Tratamiento (ITT).

Sin embargo, dado que hay un porcentaje de diadas que habiendo sido seleccionadas al grupo tratamiento, no reciben el tratamiento, para calcular el *Local Average Treatment Effect (LATE)*, se utiliza 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 utiliza para estimar el efecto del programa.

De esta forma se estiman las siguientes ecuaciones:

$$T_i = \alpha_0 + \alpha_1 C J_i + \varepsilon_i \quad (2)$$

Donde T_i es una variable que toma el valor 1 si es que la diada i fue efectivamente tratada y cero de lo contrario.

Y luego se estima la siguiente ecuación en base a los valores estimados de la ecuación (2):

$$Y_i = \beta_0 + \beta_1^{LATE} \hat{T}_i + \beta_n X_{ni} + \mu_i \quad (3)$$

Donde $\hat{T}_i$ es la probabilidad estimada de ser tratado por el programa y β_1^{LATE} corresponde al efecto sobre los *compliers*, es decir, el efecto sobre aquellas diadas que reciben el programa dado que fueron asignadas al grupo tratamiento.

4.4 Plan de análisis LATE

La evaluación del LATE se llevará a cabo para todas las variables de resultado, es decir para los outcomes primario, secundarios e intermedios.

Para definir la dosis necesaria para ser considerado efectivamente tratado, se utilizarán los criterios previamente establecidos por Fundación Infancia Primero. Según estos criterios, una diada participa de una sesión solo si descarga al menos un material de apoyo audiovisual y participa en al menos una de las actividades de interacción (ya sea participando en el foro comunitario o asistiendo al encuentro de forma presencial u online). Basándose en esta información, se define que la diada es efectivamente tratada por el programa si participa en al menos 5 sesiones.

De todas formas, se realizarán análisis de sensibilidad para identificar si existen cambios en el efecto del programa a medida que se hacen más o menos exigentes las condiciones de participación.

4.5 Análisis adicionales

4.5.1 Efectos de interacción

Una de las particularidades del programa Crecer Jugando es que en ciertas comunas la intervención es realizada en modalidad *online*. A partir de esta realidad cabe preguntarse si la intervención presencial genera mayores o menores impactos que el tratamiento remoto.

Para medir dicha realidad, se propone agregar a la ecuación (1) una variable de interacción que dé cuenta del mayor o menor efecto que puede tener el tratamiento cuando el individuo es tratado de forma presencial:

$$Y_i = \beta_0 + \beta_1 CJ_i + \beta_2 CJ_i * P_i + \beta_n X_{ni} + \mu_i \quad (4)$$

En la ecuación (4) la variable P_i toma el valor 1 si el individuo recibe tratamiento presencial y cero si lo recibe online, por lo que β_1 representa el efecto del tratamiento para el grupo online, y β_2 captura el efecto adicional (si es que este fuera positivo) que entregaría la intervención presencial con respecto a la online. De esta forma el efecto del tratamiento presencial sería $\beta_1 + \beta_2$ y el efecto del tratamiento online β_1 .

Un segundo experimento consiste en tratar ambas investigaciones como tratamientos distintos, es decir, por un lado, está la intervención de manera presencial y por otro lado de manera online.

Para medir el impacto diferenciado de dichas intervenciones proponemos la siguiente ecuación:

$$Y_i = \beta_0 + \beta_1 CJ_i^P + \beta_2 CJ_i^O + \beta_n X_{ni} + \mu_i \quad (5)$$

Donde CJ_i^P es una variable que toma el valor 1 si la diada recibe el tratamiento presencial y cero si no y CJ_i^O otra variable que toma el valor 1 si la diada recibe la intervención online y cero de lo contrario.

Es importante resaltar que, en esta situación, se produce un aumento en el efecto mínimo detectable. Este aumento se debe a la disminución en el tamaño del grupo de tratamiento, ya que ahora se compone exclusivamente de aquellos que reciben la intervención de manera presencial o de manera online por separado, a diferencia de los casos anteriores donde se aplicaba de forma conjunta. Esto

implica que se necesita un mayor tamaño de efecto para alcanzar significancia estadística debido al menor número de observaciones en cada subgrupo.

4.5.2 Análisis de robustez

Una posible fuente de sesgo es la existencia de díadas que abandonan la investigación. Como se mencionó anteriormente, esta deserción no causaría problemas de endogeneidad si es aleatoria y no está relacionada con las características de los cuidadores o niños evaluados.

Para evaluar la existencia de este sesgo, se utilizará una regresión logística para determinar si existe alguna relación entre la probabilidad de deserción y las características observables recopiladas en la línea base. Cabe mencionar que este análisis solo considerará a aquellas díadas que abandonan la participación después de responder la encuesta o entrevista inicial, dejando fuera a aquellas lo hacen previo a la aplicación de estos instrumentos.

Si bien una evaluación experimental no requiere de una encuesta de línea base, en este caso de igual forma se aplicará a ambos grupos antes del tratamiento. Esta información nos permitirá corroborar el balance entre los grupos tratamiento y control donde, a partir de un análisis de diferencia de medias, se evaluará si ambos grupos tienen diferencias significativas en alguna variable, y en el caso de que eso ocurra, se podrá controlar posteriormente por esa característica para evitar sesgos en la estimación.

Adicionalmente, se realizará un análisis de diferencia de medias entre los grupos que completan su proceso de investigación, para identificar si existe algún desbalance en la muestra debido a la deserción. En las secciones venideras se realizarán estos análisis a partir de la información recopilada a la fecha.

Cuadro 2. Resumen de diseño de evaluación

Diseño experimental, incluyendo el número de brazos		Experimento aleatorio simple, de dos brazos
Unidad de aleatorización		Díadas (cuidador/a + niño/a)
Outcome primario finales	variable	- Calidad del vínculo entre niño/a y cuidador/a - Desarrollo de habilidades lingüísticas, motoras, de resolución de problemas y sociales del niño o niña
	medida (instrumento, escala, fuente)	Calidad del vínculo entre niño/a y cuidador/a Instrumento: PICCOLO (The Parenting Interactions with Children: Checklist of Observations Linked to Outcomes) Escala: Listado de verificación de 29 comportamientos observables. Se puntúa con entre 0 y 2 en cada ítem y se obtienen 4 sub-escalas (Afecto, Responsividad, Aliento o Ánimo y Enseñanza) con 7 ítem cada una, y una escala general. Fuente: Videos de 10 minutos del cuidador/a interactuando con niño/a. Desarrollo de habilidades lingüísticas, motoras, de resolución de problemas y sociales Instrumento: ASQ-3 (Ages and Stages Questionnaires, Third Edition) cuenta con un total de 21 cuestionarios, cada uno diseñado para evaluar el desarrollo de los niños en rangos de edad específicos. Cada cuestionario está dirigido a una ventana de edad con los siguientes intervalos por meses: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, 33, 36, 42, 48, 54 y 60. Se centra en cinco áreas del desarrollo: comunicación, motricidad gruesa, motricidad fina, resolución de problemas, y habilidades personales y sociales

		● Escala: Cada cuestionario del ASQ-3 está compuesto por preguntas en cada una de las cinco áreas del desarrollo. El puntaje número de preguntas por cuestionario, pero en su gran mayoría es de 6 por cada área de desarrollo. Cada pregunta es evaluada con 1 (realiza la acción), 0,5 (realiza la acción, pero no de manera constante) y 0 (no realiza la acción). ● Fuente: Cuestionario autoaplicado a cuidadores/as
Outcome primario intermedios	variable(s)	● Frecuencia de interacciones con niño/a ● Habilidades de buen trato ● Habilidades parentales (E2P)
	medida (instrumento, escala, fuente)	Frecuencia de interacciones con niño/a ● Instrumento: Encuesta inicial/final Infancia Primero ● Escala: se evalúan a través de 7 ítem la frecuencia con que el/la cuidador/a realizó actividades con el/la niño/a (Ninguna vez en la última semana, 1 a 2 días de la última semana, 3 a 4 días de la última semana, 5 a 6 días en la última semana o todos los días de la última semana) ● Fuente: Cuestionario autoaplicado Habilidades de buen trato ● Instrumento: Encuesta inicial/final Infancia Primero ● Escala: se evalúa a través de 5 ítems con escala Likert de 5 categorías de nivel de acuerdo ● Fuente: Cuestionario autoaplicado Habilidades parentales

		● Instrumento: E2P (Escala de parentalidad positiva) ● Escala: 54 ítems, con cuatro puntos para cada ítem, entre 1 (Casi Nunca) y 4 (Siempre). ● Fuente: Cuestionario autoaplicado
Outcome(s) secundario(s)	variable(s)	● Apoyo social cuidadores ● Estrés parental
	medida (instrumento, escala, fuente)	Apoyo social cuidadores ● Instrumento: Encuesta inicial/final Infancia Primero ● Escala: se evalúa a través de 3 ítems con escala Likert de 5 categorías de nivel de acuerdo ● Fuente: cuestionario autoaplicado Estrés parental ● Instrumento: El PSI-4-SF (Parenting Stress Index, Fourth Edition, Short Form) herramienta utilizada para evaluar el nivel de estrés en los padres en relación con la crianza de sus hijos. Este instrumento tiene tres subescalas principales: Estrés Parental (Parental Distress - PD), Interacción Disfuncional Padre-Hijo (Parent-Child Dysfunctional Interaction - P-CDI) y Características Dificiles del Niño (Difficult Child - DC). El PSI-4-SF también proporciona un Índice Total de Estrés ● Escala: se evalúan a través de 36 ítems a través de una escala Likert de 5 categorías de nivel de acuerdo. Cada ítem se califica en función de estas categorías, y la puntuación se suma para cada subescala. El índice total de estrés se calcula sumando las puntuaciones de las tres subescalas.

		● Fuente: Encuesta autoaplicada
Linea base resultados intermedios	variable	● Frecuencia de interacciones con niño/a ● Habilidades de buen trato
	medida (instrumento, escala, fuente)	Frecuencia de interacciones con niño/a ● Instrumento: Encuesta inicial Infancia Primero ● Escala: se evalúan a través de 7 ítem la frecuencia con que el/la cuidador/a realizó actividades con el/la niño/a (Ninguna vez en la última semana, 1 a 2 días de la última semana, 3 a 4 días de la última semana, 5 a 6 días en la última semana o todos los días de la última semana) ● Fuente: Cuestionario autoaplicado Habilidades de buen trato ● Instrumento: Encuesta inicial/final Infancia Primero ● Escala: se evalúa a través de 5 ítems con escala Likert de 5 categorías de nivel de acuerdo ● Fuente: Cuestionario autoaplicado
Línea base para el outcome secundario	variable	Apoyo social cuidadores
	medida(s) (instrumento, escala, fuente)	Apoyo social cuidadores ● Instrumento: Encuesta inicial/final Infancia Primero ● Escala: se evalúa a través de 3 ítems con escala Likert de 5 categorías de nivel de acuerdo ● Fuente: cuestionario autoaplicado
Variables de	Variables	Nivel educacional cuidador/a

caracterización y control		Género cuidador/a Género niño/a Edad cuidador/a Edad niño/a Nacionalidad Relación diada (padre, madre, tío/a, etc.) ¿Niño/a asiste a establecimiento educacional? ¿Niño/a tiene alguna discapacidad? Situación laboral cuidador/a Número de hijo/a (1ro, 2do, 3ro, etc.)
	medida(s) (instrumento, escala, fuente)	● Instrumento: Entrevista inicial Infancia Primero de caracterización que recoge información sociodemográfica y de caracterización de la diada participante ● Fuente: Entrevista estructurada

Fuente: elaboración propia

5. 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):

$$MDES = t_{\alpha} + t_{1-\kappa} \sqrt{\frac{\sigma^2}{P(1-P)N}} \quad (5)$$

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 considera 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 donde se desea minimizar el error Tipo 1 (significancia del 5%) y garantizar una probabilidad alta de detectar un efecto verdadero (poder del 80%)

Con respecto al tamaño de la muestra, dado el presupuesto de la investigación se calcula un tamaño máximo de 460. A partir de (5) se obtiene un MDES de 0,23 DE.

Fundación Infancia Primero calculó inicialmente una tasa de atrición del 15%, lo que significaba que aproximadamente el 15% de los individuos asignados a los grupos de control o tratamiento no podrían ser evaluados con los instrumentos previstos. Para compensar esta tasa de atrición proyectada, se planeó comenzar con una muestra inicial de 540 participantes. Esto garantizaría que, después de la atrición, queden aproximadamente 460 familias, manteniendo un MDES de 0,23 DE.

Cabe destacar que para evitar sesgos que puedan comprometer la validez interna de la investigación, es crucial que se cumpla con el supuesto de que la atrición es independiente de la asignación del tratamiento. En otras palabras, la pérdida de contacto no debe estar relacionada con ser seleccionado en alguno de los dos grupos.

Tomando en cuenta un estudio de características similares como fue la Evaluación de impacto del Programa “Nadie es Perfecto” (World Bank, 2017), donde se encontraron efectos de entre 0.2 y 0.4 DE en desarrollo del lenguaje, un efecto mínimo detectable de 0,23 DE pareciera suficientemente bajo. Más aun, considerando que el programa Crecer Jugando es más intensivo y con mayor adhesión (se espera que sea mayor al 90%)⁵, y por lo tanto con mayores efectos esperados que el programa “Nadie es perfecto”.

6. Universo y participantes

Si bien el programa se implementa en 14 comunas del país, la investigación se realizará en las siguientes comunas: El Bosque, Melipilla, Quinta Normal, Limache, Concepción, Maule, Monte Patria y Rio Negro, que corresponden a las comunas que cuentan con equipos estables y capacitados en los territorios según el modelo del programa.

El protocolo inicial de esta evaluación propone medir una muestra de 460 diadas. Para lograr ese tamaño muestral, la evaluación contempla la recolección de información en al menos 3 periodos durante el año. Al comienzo de cada periodo se aplica una entrevista y encuesta de entrada tanto a controles como tratamiento, y una vez realizada la intervención, se aplican distintos instrumentos que permitirán medir la efectividad de la intervención en diferentes variables de interés.

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

⁵ Cuando hablamos de adhesión nos referimos al porcentaje de diadas que, habiendo sido asignadas al grupo tratamiento, son efectivamente tratadas. Por lo tanto, la no adhesión no es lo mismo que la atrición, dado que este último hace referencia al porcentaje de personas que, habiendo sido seleccionadas para la investigación (control o tratamiento), por alguna razón no es posible aplicarles los instrumentos.

Grupo tratamiento:

1. A partir de la semana 9 del tratamiento, se contacta telefónicamente a las diadas para invitarlas a participar del estudio y agendar la aplicación del instrumento PICCOLO (las diadas están al tanto que van a ser contactadas por las investigadoras gracias a la información entregada por los equipos implementadores).
2. Finalizado el llamado telefónico, se envía a través de whatsapp el link del cuestionario 1 (E2P y PSI-4) y el consentimiento informado para participar de la evaluación.
3. Se aplica el instrumento PICCOLO en la fecha y horario convenido y, finalizada su aplicación, se envía link del cuestionario II (ASQ-3 y encuesta de salida) a través de whatsapp.

Grupo Control:

1. Se les contacta telefónicamente para indicarles que quedaron en la lista de espera para recibir el tratamiento y se les informa la fecha aproximada de inicio de su participación.
2. Se les invita a participar de la evaluación y, a quienes acceden, se les aplica la entrevista y encuesta inicial.
3. Una vez contactadas todas las diadas se les envía un correo electrónico con el consentimiento informado para participar de la evaluación.
4. A partir de la semana 7 de la intervención del grupo tratamiento, se contacta telefónicamente a este grupo para aplicar el cuestionario I (E2P y PSI-4) y agendar el instrumento PICCOLO. Se les envía link del cuestionario I a través de whatsapp para su autoaplicación.
5. Se aplica el instrumento PICCOLO en la fecha y horario convenido y, finalizada su aplicación, se envía link del cuestionario II (ASQ-3 y encuesta de salida) vía whatsapp.

7. Análisis de costo efectividad

La evaluación de costo-efectividad conlleva las siguientes etapas:

1. Definir los costos de la intervención (costos de inversión y costos de operación) y los costos de un programa alternativo similar (ej: programa Nadie es Perfecto que ya fue evaluado a través de una evaluación de impacto). En ambos casos los costos serán ajustados a su valor social en base al Sistema Nacional de Inversiones (SIN).
La información de los costos del programa Crecer Jugando será entregada por la fundación Infancia Primero. Por su parte, los costos del programa Nadie es Perfecto (y otros programas de carácter similar) son públicos por tratarse de una iniciativa implementada por el Ministerio de Salud de Chile.
2. Determinar el valor presente de los costos considerando una tasa de descuento del 6% (SIN, 2023) y un horizonte temporal de 5 años (dependiendo de la naturaleza de los costos)
3. En base a los resultados de la evaluación de impacto y el cálculo del valor presente de los costos, se vislumbran los siguiente cuatro posibles escenarios:

- i. El programa Crecer Jugando tiene un mayor efecto que sus alternativas y un menor costo.
- ii. El programa Crecer Jugando tiene un menor efecto que sus alternativas y un mayor costo.
- iii. El programa Crecer Jugando tiene un mayor efecto que sus alternativas y un mayor costo.
- iv. El programa Crecer Jugando tiene un menor efecto que sus alternativas y un menor costo.

En las alternativas (i) y (ii) no es necesario llevar a cabo mayores análisis, porque en el caso (i) se concluiría que el programa Crecer Jugando claramente es más costo-efectivo que sus alternativas y en el caso (ii) claramente menos costo efectivo.

En el caso que nos encontremos en las situaciones (iii) y (iv), se analizará lo que se llama *costo efectividad incremental*:

$$ICER = \frac{C_{crecer\ jugando} - C_{programa\ alternativo}}{E_{crecer\ jugando} - E_{programa\ alternativo}}$$

Donde C es el costo y E el efecto del programa. En base al cálculo del ICER se evaluará si este está sobre o bajo un umbral de costo efectividad que será definida junto a la contraparte en base a sus expectativas sobre el programa y a evidencia nacional e internacional de programas de naturaleza similar.

8. Análisis de fidelidad

La evaluación de procesos se realiza durante la implementación de un programa o iniciativa para conocer la fidelidad de la ejecución, es decir, si se desarrolla de acuerdo a lo diseñado (Iválua, 2009). Estas buscan proveer información relevante respecto a la concreción del programa según lo programado, permitiendo identificar aspectos críticos y brechas del proceso a nivel de actividades, cronograma, recursos, responsables y población objetivo. Este tipo de enfoque es pertinente en etapas iniciales de un programa, permitiendo optimizar la ejecución, dado que detecta problemas en la implementación y sugiere soluciones pertinentes que permitan avanzar hacia los resultados esperados.

En sus casi 9 años de implementación, Crecer Jugando ha realizado, en base a la experiencia y la realización de investigaciones, una serie de evaluaciones de procesos que le han permitido realizar los ajustes pertinentes a su diseño, iterando entre distintas alternativas de intervención logrando la combinación más eficiente para lograr los resultados esperados. Estas evaluaciones han sido puestas a disposición de SUMMA.

En este contexto, se propone la realización de una evaluación de fidelidad (y no de procesos) con dos objetivos. Un primer objetivo, y quizás el más evidente, identificar si la implementación en las 9 comunas en las que se realizará la evaluación se desarrolló tal como fue planificada. Un segundo objetivo de estos indicadores es la posibilidad que nos entregan para detectar el efecto de la intervención sobre las familias efectivamente tratadas. Como ya se mencionó en la sección 4, existen dos tipos de mediciones, el ITT y el LATE.

Al término de la intervención total y antes de la consolidación de las bases de datos para el análisis, se realizará en análisis de procesos, en base a información secundaria, la que es recolectada como parte de los procesos de monitoreo permanentes del programa Crecer Jugando.

La participación en el programa se puede concretar a través de dos criterios. En primer lugar, se considera la asistencia a las sesiones presenciales u online, donde las diadas realizan actividades en conjunto con los implementadores de la Fundación. En segundo lugar, se define un criterio de participación basado en la descarga de los materiales enviados por la Fundación a las diadas que forman parte del grupo.

9. Consideraciones éticas

Para la investigación se cuenta con la aprobación del Comité de Ética de la Universidad Alberto Hurtado. Este comité está facultado para:

1. Velar por el respeto de los derechos de los sujetos de investigación relativos a la ponderación riesgo beneficio, confidencialidad-anonimato, y libertad de decisión de los sujetos investigados
2. Resguardar a los investigadores y a la Universidad en cuanto a la realización de estudios que, cumpliendo con los estándares éticos pertinentes, aseguren un buen desarrollo de los mismos
3. Solicitar modificaciones o enmiendas a la investigación sobre los aspectos éticos que, a juicio del comité, así lo requiera.
4. Rechazar el proyecto de investigación cuando este no cumpla con los estándares éticos elegibles. Las disposiciones del Comité son vinculantes y finales, no procediendo en consecuencia contra ellas ningún recurso o revisión ulterior.
5. Ordenar la suspensión o terminar la ejecución del estudio hasta que no existan resguardos suficientes del cumplimiento de los estándares éticos pertinentes, previa reunión con los responsables del estudio.
6. Otorgar certificación ética para las instancias que lo exigen, tales como los recursos de proyectos de fondo público y fondo interno de investigación.
7. Certificar los proyectos que requieren de un título de aprobación ética y visado del formulario de consentimiento informado por parte del Comité de Ética de la Universidad, según el procedimiento detallado que se encuentra en el reglamento del Comité de Evaluación Ética de la UAH

En cuanto a las consideraciones éticas particulares del proyecto. Para resguardar la **seguridad de los participantes** en el levantamiento de información, se consideran 2 estrategias diferenciadas:

1. Conducta apropiada de los y las investigadoras, en concreto cumplir con los siguientes puntos:
 - a. **Comportamiento.** El/la profesional tendrá un comportamiento responsable y ético al interior (y exterior) de los espacios en que se realice el levantamiento de

información, donde se velará por proteger en todo momento a los participantes, anteponiendo su bienestar a los intereses de la investigación. El/la investigador/a será respetuoso(a) con los espacios, materiales de trabajo y personas que se encuentren en el espacio en que se realiza el levantamiento.

- b. **Imagen.** El/la profesional siempre mostrará una imagen profesional y apropiada. Consecuentemente, la vestimenta y arreglo personal será coherente con la función que desempeña cada colaborador(a) en el entorno que se desenvuelva.
 - c. **Lenguaje.** El profesional también tendrá un lenguaje verbal y corporal adecuado.
 - d. **Respeto por las normas.** El profesional mostrará una actitud de respeto por las normas implícitas de conducta del espacio institucional en que se realiza la entrevista.
 - e. **Confidencialidad.** El profesional protegerá la confidencialidad de los participantes del estudio. Para ello cada ejercicio investigativo contará con el consentimiento, previo y formal, de los participantes, quienes reconocen los objetivos del estudio y en qué consiste su participación.
2. Protocolo de seguridad de los participantes al momento de realizar el levantamiento, en concreto cumplir con los siguientes puntos
- a. **Consentimiento informado.** Se leerá en voz alta a los/as cuidadores/as participantes de cada actividad el Consentimiento Informado. En éste, se explicitará que no existen beneficios personales ni riesgos asociados a su participación. También, se indicará que la información levantada es de carácter confidencial, y la voluntariedad de la participación.
 - b. **Seguridad y confidencialidad de la información.** El participante será informado, previo al inicio de la actividad, mediante el consentimiento informado, de la confidencialidad y del procedimiento general de resguardo de la información que se levantará por medio de la actividad.
 - c. **Seguridad del espacio.** Las entrevistas se realizarán en un lugar seguro, en términos de infraestructura, y también asegurando un espacio privado, que permita al participante expresarse con tranquilidad y seguridad.
 - d. **Contingencias durante el desarrollo de la instancia.** Ante contingencias o situaciones que se presenten el levantamiento y que afecten el bienestar de los participantes (por ejemplo, urgencias de tipo profesional, sensibilidad expresa respecto del abordaje de un tema, o cualquier otro evento que dificulte a la persona para seguir con la instancia), el investigador ofrecerá de inmediato una pausa o, según sea el caso, terminar la instancia. Se evaluará la pertinencia de continuar en otro momento, o bien reemplazar al informante.

Para la **protección de la identidad de las personas** que participen de las actividades de levantamiento de información, se consideran los siguientes procedimientos:

- a. Se dará el consentimiento informado a todos los participantes donde se especificará que la identidad de los informantes será protegida en todo el proceso.
- b. La información será inicialmente resguardada en un archivo videos (PICCOLO) o base de datos, en cuyo nombre no se identificará el nombre del o la informante, sino información como cargo (si aplica), y el número correlativo a la comuna donde se desempeña.
- c. A los videos de interacción entre el/la cuidador/a y el niño o la niña, solo tendrán acceso un equipo de investigadores, por medio de Google Drive. Este equipo previamente deberá manifestar activamente y de manera expresa su compromiso de no compartir los datos incluidos en el archivo de audio.
- d. La generación de los reportes permitirá describir y analizar a los participantes forma conjunta, pero no individualizar a los informantes.
- e. Al finalizar la consultoría, el jefe de proyecto recopilará toda la información levantada (videos, audios de entrevistas, transcripciones, etc.), dispuesto tanto en los computadores utilizados como el servidor externo y se encargará de borrar todos los archivos confidenciales. Se guardará un respaldo hasta 6 meses después de la presentación de resultados.

Por su parte, se consideran los siguientes **principios generales para los profesionales** que levantarán la información y quiénes la analizarán:

- a. Los/las profesionales deberán adoptar siempre métodos de trabajo seguros y tomar las debidas precauciones para salvaguardar la seguridad y para protegerse de pérdidas de información, elaborando estrategias adecuadas de obtención de copias de seguridad.
- b. Los expedientes y registros originales se pueden copiar o examinar en la propia oficina o el lugar de trabajo del analista, caso en el cual se debe tomar toda precaución razonable para evitar la posibilidad de pérdida, robo o acceso por personas no autorizadas.
- c. El acceso a los datos solo estará disponible para el equipo de trabajo, y, de manera anonimizada, tal como lo señala el procedimiento anterior.
- d. Toda la información que no sea de conocimiento público deberá ser tratada como información confidencial. En ese sentido, dicha información no podrá ser divulgada ni revelada salvo que se cuente con autorización de la contraparte.
- e. Toda actividad de levantamiento de información contará con la aprobación previa de las partes involucradas, quienes acuerdan participar en la investigación y consienten el utilizar los datos obtenidos con fines de cumplir los objetivos estipulados en el estudio.
- f. La investigación deberá atenerse necesariamente a la normativa legal vigente de cada territorio en términos de protección de datos personales.

10. Equipo

El equipo de evaluación corresponde a Focus, organización con más de 30 años experiencia en el desarrollo de estudios, evaluaciones, programas y asesorías para organizaciones del sector público y privado. Los objetivos y requerimientos de esta evaluación se alinean a la misión de Focus respecto a “desarrollar y fortalecer políticas, programas y estudios innovadores y transformadores en educación, niñez, desarrollo social y sostenible, que impacten positivamente en las personas y en la sociedad, especialmente en aquellos en situación de vulnerabilidad”. El equipo a cargo de la ejecución del estudio está compuesto por profesionales sociólogos, economistas y psicólogos con amplia experiencia en evaluación de impacto e implementación, aplicación de instrumentos estandarizados, observación y coordinación de trabajos de campo de gran envergadura. El equipo está compuesto por:

Jefa de proyecto. Isidora Fuenzalida: Socióloga de la Universidad Católica, Magíster en Políticas Públicas de la Universidad de Chile. Experiencia conl foco en el diseño, evaluación y fortalecimiento de programas sociales y políticas públicas.

Coordinador de terreno. Juan Pablo Madrid. Sociólogo Universidad Alberto Hurtado. Magíster en Gobierno y Asuntos Públicos FLACSO-México. Experiencia en el diseño, implementación y evaluación de programas sociales.

Coordinador de impacto. Benjamín Coloma: Economista de la Universidad Católica, Magíster de Economía de la Universidad Católica. Especialista en la evaluación de Políticas Públicas.

Coordinador de análisis. Benjamín Nielsen: Psicólogo de la Universidad Católica, Magíster en Ciencias Sociales de la Universidad de Amsterdam. Especialista en investigación social con métodos mixtos avanzados.

Equipo de análisis Piccolo: investigadore/as con experiencia y certificación en análisis y codificación de Piccolo:

- Camila Pinto (Psicóloga clínica e investigadora del Centro CUIDA)
- Catalina Undurraga (Magister en psicoterapia, candidata a doctor en psicoterapia)
- Joaquín Grez: Psicólogo de la Universidad Católica, Magíster en Psicología Educacional de la Universidad Católica.
- Consuelo Moreno: Psicóloga de la Universidad Católica. Magister en Psicología de la Educación de la Universidad de Barcelona

IV. Informe estadístico y estado de avance de la evaluación

1. Aleatorización y terreno

A continuación, 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 diadas, 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
El Bosque 1	90	21	21	12	12	12	12
El Bosque 2	50	0	0	15	15	10	10
Concepción P	60	18	18	0	0	12	12
Concepción O	90	21	21	12	12	12	12
Quinta Normal	52	10	10	8	8	8	8
Maule	46	0	0	15	15	8	8
Melipilla	54	15	15	12	12	0	0
Río Negro	24	12	12	0	0	0	0
Monte Patria	40	9	9	11	11	0	0
Limache	36	0	0	10	10	8	8
Total	542	106	106	95	95	70	70

A lo largo de la investigación se han identificado dos momentos en los que se produce 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 directamente no querer participar del programa (ya sea del grupo control o tratamiento). Por último, está el grupo de cuidadores/as a los que no fue posible contactar.

El segundo momento donde se observa una nueva atrición de la muestra participante, es durante la aplicación de los instrumentos finales de la investigación (PICCOLO, ASQ-3, E2P y PSI-4). A esta atrición la hemos denominado Atrición Etapa 2. Más adelante se identificarán en detalle las razones por las que los/as cuidadores/as desisten de participar de este proceso (Ver Cuadro 4)

En la Tabla 2 se presenta la atrición de la Etapa 1 en las comunas que ya terminaron el proceso de levantamiento de información, es decir, aquellas donde el grupo tratamiento ya ha recibido la intervención completa y donde, adicionalmente, se han contactado a todos los cuidadores y

cuidadoras para aplicar los instrumentos finales⁶. Estas comunas y sus respectivos ciclos son: El Bosque Ciclo 1, Concepción Presencial Ciclo 1, Concepción Online Ciclo 1 y Quinta Normal Ciclo 1.

En primer lugar, se observan niveles de atrición disímiles tanto entre las comunas, como también entre los grupos tratamiento y control. Por un lado, en el grupo tratamiento, la comuna de Concepción, en su versión online, presenta niveles de atrición del 46%, valor muy distinto a la comuna de El Bosque que alcanza una atrición de apenas el 13%. 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 la comuna de El Bosque, tiene un nivel de atrición que supera el 50%, mientras que, al contrario de lo que ocurre en el grupo tratamiento, la comuna de Concepción Online, presenta la atrición más baja, con un 14%. Esta disminución de la atrición de las comunas respecto al Bosque 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. Este protocolo ajustado fue el que se implementó en las comunas venideras.

Tabla 2. Atrición Etapa 1

	Encuesta y entrevista completada			Atrición Etapa 1
	No	Sí	Total	
Tratamiento	32	75	107	30%
El Bosque	3	21	24	13%
Concepción P	6	20	26	23%
Concepción O	21	25	46	46%
Quinta Normal	2	9	11	18%
Control	26	76	102	25%
El Bosque	13	10	23	57%
Concepción P	5	19	24	21%
Concepción O	6	38	44	14%
Quinta Normal	2	9	11	18%
Total	58	151	209⁷	28%

Fuente: elaboración propia

Las diferencias en los niveles de atrición observados no son un problema per se, sin embargo, se deberá analizar si la atrición de quienes responden la encuesta y entrevista genera un desbalance entre los grupos de tratamiento y control en lo que respecta a características que pudiesen afectar el resultado de los instrumentos de medición final. Dicho de otro modo, se deberá analizar si la atrición pudiese generar sesgos de selección en los resultados de la intervención. Esto será revisado en detalle en la sección 3 de este informe.

⁶ La Tabla 2 considera aquellas comunas donde finalizó el levantamiento de información al momento de elaborar este informe. Los avances asociados al resto de las comunas son reportados mensualmente en las reuniones tripartitas entre Fundación infancia Primero, SUMMA y Focus.

⁷ El número total de aleatorizados es mayor al inicialmente proyectado debido, principalmente, al alto número de díadas inscritas en el ciclo 1 de Concepción Online.

En cuanto a la atrición en la Etapa 2, en la Tabla 3 se observan respuestas distintas entre los grupos control y tratamiento. Por una parte, en el grupo control se observa una atrición del 29%, muy distinta a la observada en el grupo tratamiento, que alcanza el 67%. Al respecto, cabe señalar que inicialmente se proyectó una atrición mayor por parte de las diadas del grupo de control, por lo que se definió como incentivo la entrega de una Gift Card de \$20.000 pesos (a canjear en diferentes tiendas nacionales) para quienes formaran parte de la evaluación. En el caso de las diadas del grupo tratamiento se proyectó una fidelización mayor por el hecho de haber participado del programa y, por lo mismo, se estableció un incentivo menor: el sorteo de una Gift Card de \$30.000 pesos por ciclo de implementación entre todos los participantes del estudio⁸.

Cabe mencionar que en la Tabla 3 se presentan aquellas diadas que completan la investigación, es decir que responden la totalidad los cuatro instrumentos (PICCOLO, ASQ-3, E2P y PSI-4), dejando afuera a aquellas diadas que pueden haber respondido solo una parte de ellas. Más adelante, en los análisis finales, se considerará como muestra a las diadas que hayan respondido el instrumento correspondiente al análisis. Por lo tanto, el tamaño total de la muestra puede variar según el *outcome* evaluado, en función de la cantidad de respuestas obtenidas para cada uno.

Tabla 3. Atrición Etapa 2

	Completa la investigación		Total	Atrición Etapa 2
	No	Sí		
Tratamiento	50	25	75	67%
El Bosque	14	7	21	67%
Concepción P	12	8	20	60%
Concepción O	18	7	25	72%
Quinta Normal	6	3	9	67%
Control	22	54	76	29%
El Bosque	3	7	10	30%
Concepción P	5	14	19	26%
Concepción O	13	25	38	34%
Quinta Normal	1	8	9	11%
Total	72	79	151	48%

Fuente: elaboración propia

Nuevamente, dado que los niveles de atrición entre ambos grupos son significativamente distintos, es necesario analizar si estas diferencias podrían generar sesgos en la estimación del efecto del programa. Más adelante se realizarán algunos ejercicios para identificar en qué medida la atrición diferenciada de quienes terminan las actividades consideradas en la evaluación genera un desbalance entre ambos grupos. Esto será revisado en detalle en la sección 3 de este informe.

En el Cuadro 3 se presenta un resumen de las etapas de aleatorización. Al respecto, se puede observar que en una primera Etapa se produce una pérdida de 58 diadas seguido de una pérdida de

⁸ Actualmente se está analizando con SUMMA e Infancia Primero la posibilidad de contar con recursos adicionales para entregar una Gift Card de \$20.000 tanto a las diadas del grupo de control como del grupo de tratamiento.

72, lo que da como resultado un total de 130 diadas que se pierden en el proceso, es decir, un porcentaje de atrición del 62%

Cuadro 3. Resumen Etapas de Atrición

Diadas Aleatorizadas	Completan la entrevista y la encuesta	Completan investigación
209	151	79
	Atrición Etapa 1	Atrición Etapa 2
	58 (28%)	72 (48%)
	Atrición Total	
	130 (62%)	

Fuente: elaboración propia

El equipo de investigación ha ido monitoreando y categorizando las razones de desistimiento por parte de los/as cuidadores contactados/as, a partir de las cuales se han ido definiendo, en conjunto con el equipo de Infancia Primero, una serie acciones que buscan disminuir la atrición. Estas acciones también han sido presentadas y discutidas en las reuniones tripartitas mensuales con el equipo de SUMMA. Las acciones que se han ido poniendo en práctica se resumen en el Cuadro 4.

Cuadro 4. Acciones para aumentar la tasa de respuesta

Categorías	Porcentaje	Razones	Acciones
Participación parcial	9,6%	En proceso	Protocolo en ejecución
		Resistencia a ser grabado	Ajustes en el discurso, reforzamiento de confidencialidad y seguridad de la información
No manifiesta interés o desiste de participar	47,2%	Desinformación sobre el programa (grupos tratamiento y control)	Derivación a los equipos comunales respectivos (planilla de contacto compartida por IP)
		Desinformación sobre la evaluación en grupo tratamiento (baja vinculación con el programa, participación incompleta)	Documento "Preguntas y respuestas sobre la evaluación" para equipos comunales
			Infancia Primero refuerza discurso con equipos territoriales Estrategias de rescate con equipo Focus complementario
			Focus se suma a la última sesión del programa (modalidad online)
			Focus se reúne con equipo territorial en Concepción
			Focus comenzará a sumarse a primera sesión del programa (A partir de segundo ciclo de El Bosque)

Categorías	Porcentaje	Razones	Acciones
		Razones personales (tiempo, trabajo, institucionalización)	No aplica
Deja de contestar	24,8%	Hipótesis A: Desistimiento indirecto	Estrategia con equipo Focus complementario Búsqueda de contacto a través de RRSS Contacto con supervisoras/es o facilitadores/as para recabar otros contactos
		Hipótesis B: Extravío, pérdida o cambio de teléfono	
No es posible contactar a la persona	18,4%	Buzón de voz	
		Número no existe	
		Número equivocado	

Fuente: elaboración propia

Ante la constatación de altos porcentajes de atrición, significativamente superiores a los inicialmente previstos, se realizó un esfuerzo por parte de Infancia Primero por incrementar el número de diadas a aleatorizar, con el objetivo de lograr un número final de diadas analizadas similar a las 460 planificadas inicialmente.

A continuación, en la Tabla 4 se presenta una proyección de las diadas a aleatorizar en los ciclos 2 y 3, con lo que se espera lograr un total de 640 diadas aleatorizadas.

Tabla 4. Aleatorización proyectada

		Actual		Proyectado			
		Ciclo 1		Ciclo 2		Ciclo 3	
Comuna	Muestra	T	C	T	C	T	C
El Bosque 1	95	24	23	12	12	12	12
El Bosque 2	50			15	15	10	10
Concepción P	94	26	24			22	22
Concepción O	140	46	44			25	25
Quinta Normal	58	11	11	10	10	8	8
Maule	47	16	15			8	8
Melipilla	55	16	15	12	12		
Río Negro	25	11	14				
Monte Patria	40			9	9	11	11
Limache	36			10	10	8	8
Total	640	150	146	68	68	104	104

Fuente: elaboración propia

A continuación, se presenta la siguiente proyección en base a la actualización de diadas disponibles para aleatorizar y a los niveles de atrición reales que han existido hasta ahora. En base a esta nueva

proyección, y considerando un universo en donde ambos grupos mantienen en los ciclos 2 y 3 sus niveles actuales de atrición, en la Tabla 5 se presenta una proyección de la muestra final.

Tabla 5. Resumen proyección final de la muestra

	Tratamiento	Control
Díadas Aleatorizadas	322	318
Atrición	77%	47%
Muestra Final	74	169

Fuente: elaboración propia

2. Actualización y proyección del Efecto Mínimo Detectable

En base a los nuevos cálculos de atrición y la nueva proyección de díadas a aleatorizar, en la Tabla 6 se muestra el MDES tanto para el protocolo inicial como para la nueva proyección. En base a una atrición de 77% para el grupo control y de 47% para el grupo tratamiento, sumado a las nuevas proyecciones evidenciadas en la Tabla 4, se calcula un MDES proyectado de 0,35 DE.

Tabla 6. Efecto Mínimo Detectable

		Protocolo	Proyectado
		TOTAL	TOTAL
Tamaño mínimo del efecto detectable (MDES)		0,23	0,35
Alfa		0.05	0.05
Poder		0.8	0.8
¿Unilateral o bilateral?		unilateral	unilateral
Atrición Grupo Tratamiento		15%	77%
Atrición Grupo Control		15%	47%
Número de díadas	Tratamiento	270	322
	Control	270	318
	total	540	640

Fuente: elaboración propia

En el cálculo del MDES proyectado, existe un supuesto principal: que la atrición observada hasta el momento en el ciclo 1 se mantendrá constante en los ciclos 1 y 2. Este supuesto es bastante conservador y representa posiblemente el escenario más pesimista, dado que se han realizado ajustes al protocolo de contacto que deberían tener efectos positivos en la respuesta de las díadas y, por lo tanto, disminuir los niveles de atrición en ambos grupos.

Por último, es importante destacar que el tamaño de la muestra y el poder estadístico de este estudio fueron diseñados para evaluar efectos significativos a nivel global, es decir, no se pretende detectar efectos dentro de subgrupos específicos. Por lo tanto, nuestros resultados podrían no tener la sensibilidad necesaria para identificar efectos significativos dentro de grupos más pequeños (por ejemplo, sobre rangos de edad, género de el niño o la niña o comuna de residencia)

3. Caracterización de los participantes (preliminar a la fecha)

5.1 Análisis de la totalidad de participantes que inician el proceso de evaluación (preliminar a la fecha)

5.1.1 Balance de grupos

A continuación, se realiza un análisis de balance entre el grupo control y tratamiento. En este primer análisis se consideran las 151 díadas que completaron tanto la entrevista como la encuesta inicial.

En la Tabla 7 se observa una caracterización de niños y niñas de ambos grupos. En cuanto a las diferencias, se evidencia que la proporción de mujeres, la proporción de niños/as con alguna discapacidad y la proporción que corresponde a primer hijo/a presentan diferencias estadísticamente significativas al 90% de confianza.

Tabla 7. Balance de caracterización general niños/as

	Tratamiento	Control	Diferencia
Edad Niño/a (promedio de meses)	18,6	18,8	-0,163
Genero Niño/a (proporción mujeres)	0,427	0,566	-0,139*
Educación Niño (proporción en establecimiento educacional)	0,373	0,342	0,031
Discapacidad (proporción con alguna discapacidad)	0,013	0,079	-0,066*
Número de hijo (proporción primer hijo)	0,667	0,789	-0,123*
Número de observaciones	75	76	

Fuente: elaboración propia Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Por su parte, en la Tabla 8 se observan las diferencias en las características de los cuidadores y cuidadoras. En este caso solo se observan diferencias estadísticamente significativas al 90% de confianza en la proporción de mujeres que componen la muestra.

Tabla 8. Balance de caracterización cuidadores/as

	Tratamiento	Control	Diferencia
Edad Cuidador/a (promedio de años)	32,880	32,079	0,801
Genero Cuidador/a (proporción mujeres)	0,933	0,987	-0,054*
Relación con el niño o niña (proporción madre)	0,960	0,974	-0,014
Nacionalidad (proporción chileno/a)	0,747	0,750	-0,003
Educación (proporción con educación superior completa o incompleta)	0,627	0,737	-0,110
Estado civil o conyugal (proporción casado/a o conviviente)	0,587	0,461	0,126
Empleo (proporción con trabajo remunerado)	0,453	0,579	-0,126
Observaciones	75	76	

Fuente: elaboración propia Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.001$

Adicionalmente se realiza un análisis de la encuesta inicial como una manera de determinar si existen diferencias iniciales entre los grupos en cuatro dimensiones: frecuencia de interacción, apoyo social, habilidades de buen trato y dificultades parentales.

Para medir la frecuencia de interacción se realizan preguntas donde los cuidadores y cuidadoras deben indicar con qué frecuencia realizan una serie de actividades con sus hijos o hijas dentro de las siguientes alternativas: ningún día de la semana, 1 a 2 días a la semana, 3 a 4 días a la semana, 5 a 6 días a la semana o todos los días de la semana. Para realizar el análisis de diferencia de medias se presenta la proporción de diadas que declara realizar todos los días de la semana la actividad correspondiente con su hijo o hija

Por su parte, para el apoyo social, habilidades de buen trato y dificultades parentales, los cuidadores y cuidadoras deben indicar qué tan de acuerdo están con una serie de afirmaciones. Para realizar el análisis de diferencia de medias se indica la proporción de diadas que están de acuerdo o muy de acuerdo con la afirmación correspondiente.

En la Tabla 9 se observa que en la gran mayoría de las variables no se identifican diferencias estadísticamente significativas, a excepción de la frecuencia con la que le cantan a sus hijos/as y la frecuencia con la que habló conversó o hizo preguntas al niño o niña sobre cosas que le interesan a el/ella. En ambos casos existe una mayor frecuencia de interacción en el grupo control.

Tabla 9. Balance línea base

	Tratamiento	Control	Diferencia
Frecuencia de interacción			
Leyó con el niño o niña un cuento, libro o revista	0,080	0,132	-0,052
Le cantó o cantó con el niño o niña	0,613	0,803	-0,189**

	Tratamiento	Control	Diferencia
Jugó algún juego de interés con el niño o niña	0,640	0,658	-0,018
Habló, conversó o hizo preguntas al niño o niña sobre cosas que le interesen a él/ella	0,573	0,711	-0,137*
Facilitó y permitió el juego libre del niño o niña	0,680	0,816	-0,136*
Realizó tareas del hogar junto con el niño o niña	0,427	0,474	-0,047
Usó momentos de la vida cotidiana para estimular al niño o niña (ejemplo: cantarle durante el baño, masaje durante la muda)	0,680	0,697	-0,017
Apoyo social			
Confío en mis pares (otras madres, padres o cuidadores) y puedo aprender de ellos	0,720	0,684	0,036
Tengo a alguien en quien confío para pedirle consejos sobre cómo ser madre, padre o cuidador	0,640	0,671	-0,031
Tengo alguien cercano a quien pedir apoyo en momentos donde me he sentido estresado/a o sobrepasado/a con la crianza	0,733	0,750	-0,017
En las últimas dos semanas me he sentido solo/a o aislado/a	0,267	0,263	0,004
Habilidades de buen trato			
Logro comprender las necesidades de mi niño o niña	0,907	0,895	0,012
Tengo herramientas para manejar situaciones desafiantes en la crianza	0,773	0,789	-0,016
Respondo de manera paciente y con un buen trato cuando mi niño o niña se porta mal o se equivoca	0,853	0,829	0,024
Logro comprender las emociones de mi niño o niña y puedo ayudarlo/a a manejarlas de buena manera	0,893	0,855	0,038
Logro comprender que hay detrás de la conducta de mi niño o niña o lo que busca expresar con esa conducta	0,800	0,737	0,063
Dificultades parentales			
Atender a mi hijo/a a veces me quita más tiempo y energía de la que tengo	0,480	0,434	0,046
Me resulta difícil equilibrar diferentes responsabilidades debido a mi hijo/a	0,333	0,276	0,057
El comportamiento de mi hijo/a a menudo me resulta incómodo o estresante	0,133	0,145	-0,011
Me siento abrumado/a por la responsabilidad de ser padre/madre	0,160	0,105	0,055
La mayor fuente de estrés en mi vida es mi hijo/a	0,053	0,013	0,040
¿Cree que su niño o niña tiene un desarrollo igual al de otros niños o niñas de su edad?	0,693	0,763	-0,070
Observaciones	75	76	

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.001$

5.1.2 Cumplimiento (totalidad de participantes)

A continuación, se analiza la participación de las díadas asignadas al grupo tratamiento. En la Tabla 10 se muestra su asistencia en las sesiones del programa. Lo primero que se observa es que más del 70% de quienes responden la encuesta inicial y entrevista, asisten a lo menos a una sesión. Por su parte, se observa que un 48% asiste a 5 sesiones o más y que apenas un 5,3% completa la totalidad de sesiones.

Tabla 10. Participación en tratamiento

Número de asistencias	N	%	Acumulado
No asiste a ninguna sesión	22	29,3%	
1	7	9,3%	9,3%
2	5	6,7%	16,0%
3	4	5,3%	21,3%
4	1	1,3%	22,7%
5	4	5,3%	28,0%
6	4	5,3%	33,3%
7	7	9,3%	42,7%
8	13	17,3%	60,0%
9	4	5,3%	65,3%
10	4	5,3%	70,7%
Total	75	100,0%	100,0%

Fuente: elaboración propia

A continuación, en la Tabla 11 se observan las diferencias de asistencia en los distintos territorios. Por un lado, en la comuna de Concepción se evidencia una asistencia promedio de 6,4, lo que contrasta con el promedio de Quinta Normal, donde el número de sesiones promedio asistidas alcanza apenas 2,5.

En cuanto a la asistencia del total de la muestra, se observa que esta alcanza un 5,8 promedio por día y una mediana de 7 asistencias. Cabe destacar que en la Tabla 11 solo se consideran las días que asisten a por lo menos una sesión, dejando fuera a las 22 días que, habiendo contestado los instrumentos iniciales, no participaron en ninguna sesión del programa.

Tabla 11. Estadística descriptiva de asistencia

	Concepción	El Bosque	Quinta Normal	Total
Promedio de sesiones asistidas	6,4	5,8	2,5	5,8
Mínimo	1	1	1	1
Máximo	10	10	5	10
Mediana	7	6,5	2	7
Cantidad de observaciones	31	16	6	53

Fuente: elaboración propia

Para determinar el nivel de participación en base a este segundo criterio, en la Tabla 12 se observa la cantidad de días que descargaron por lo menos alguno de los materiales enviados de forma digital (audio, invitaciones, canciones, cuentos o videos). En este caso se evidencia que el 53% ha descargado por lo menos alguno de los cuatro tipos de materiales en todas las ocasiones en las que han sido enviados. Por su parte, solo el 2,7% no ha descargado ningún tipo de material a lo largo de todo el proceso de intervención.

Tabla 12. Descarga de material

Cantidad de veces que descarga al menos un material	N	%
No descarga ningún material	2	2,7%
1	2	2,7%
2	8	10,7%
3	2	2,7%
4	7	9,3%
6	1	1,3%
7	3	4,0%
8	4	5,3%
9	6	8,0%
10	40	53,3%
Total	75	100,0%

Fuente: elaboración propia

5.2 Análisis grupo de participantes que termina el proceso de evaluación (preliminar a la fecha)

5.2.1 Balance de grupos

Uno de los principales riesgos de la atrición es que genere desbalance entre los grupos tratamiento y control, y que esto implique sesgos en las estimaciones del efecto del programa.

A continuación, se realiza un análisis del balance entre ambos grupos, pero esta vez considerando el subgrupo de díadas que, a la fecha, han completado su proceso investigación, es decir, aquellas díadas a las que se les ha aplicado todos los instrumentos.

Cabe recordar que, debido a la naturaleza de esta intervención, el número de díadas analizadas es aún baja y, por lo tanto, es probable que los resultados cambien a medida que aumente la cantidad de observaciones. Dicho lo anterior, estos resultados son una primera aproximación y cualquier hipótesis para explicar las diferencias debe tomar en cuenta el bajo tamaño de muestra que por el momento se ha logrado analizar.

En la Tabla 13 se observan solo diferencias estadísticamente significativas al 95% de confianza en la proporción de niños/as con alguna discapacidad. Con respecto al género y la proporción que corresponde a primer hijo/a, en este subgrupo desaparecen las diferencias observadas en la Tabla 7.

Tabla 13. Balance de caracterización general niños/as (grupo que termina el proceso de evaluación)

	Tratamiento	Control	Diferencia
Edad Niño/a (promedio de meses)	17,5	19,0	-1,501
Genero Niño/a (proporción mujeres)	0,560	0,574	-0,014
Educación Niño (proporción en establecimiento educacional)	0,360	0,370	-0,010
Discapacidad (proporción con alguna discapacidad)	0,000	0,093	-0,093**
Número de hijo (proporción primer hijo)	0,720	0,778	-0,058
Observaciones	25	54	

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: * $p<0.1$ ** $p<0.05$ *** $p<0.001$

En relación a las diferencias entre ambos grupos, en lo que respecta a las características de los cuidadores y cuidadoras, en la Tabla 14 se observa que no existen diferencias estadísticamente significativas. Es interesante observar que, en este grupo que termina el proceso de evaluación, desaparecen las diferencias en la proporción de mujeres evidenciadas en la Tabla 8 tanto por un aumento de la proporción en el grupo tratamiento y una disminución del grupo control.

Tabla 14. Balance de características cuidadores/as (grupo que termina el proceso de evaluación)

	Tratamiento	Control	Diferencia
Edad Cuidador/a (promedio de años)	34,040	33,074	0,966
Genero Cuidador/a (proporción mujeres)	0,960	0,981	-0,021
Relación con el niño o niña (proporción madre)	0,960	0,963	-0,003
Nacionalidad (proporción chileno/a)	0,760	0,722	0,038
Educación (proporción con educación superior completa o incompleta)	0,600	0,704	-0,104
Estado civil o conyugal (proporción casado/a o conviviente)	0,640	0,500	0,140
Empleo (proporción con trabajo remunerado)	0,520	0,611	-0,091
Observaciones	25	54	

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: * $p<0.1$ ** $p<0.05$ *** $p<0.001$

Por otra parte, en relación al balance de la línea base se observa que, salvo en dos variables, no se observan diferencias estadísticamente significativas entre los grupos. Solo se observan diferencias en la dificultad de equilibrar diferentes responsabilidades debido a su hijo/a y a la proporción que está de acuerdo con que su hijo es su principal fuente de estrés.

Es importante señalar que los datos presentados en este apartado deben interpretarse únicamente como un primer acercamiento, dado el reducido número de díadas que, hasta la fecha, han completado todos los instrumentos.

Tabla 15. Balance línea base (grupo que termina el proceso de evaluación)

	Tratamiento	Control	Diferencia
Frecuencia de interacción			
Leyó con el niño o niña un cuento, libro o revista	0,080	0,185	-0,105
Le cantó o cantó con el niño o niña	0,680	0,833	-0,153
Jugó algún juego de interés con el niño o niña	0,640	0,704	-0,064
Habló, conversó o hizo preguntas al niño o niña sobre cosas que le interesen a él/ella	0,560	0,722	-0,162
Facilitó y permitió el juego libre del niño o niña	0,640	0,796	-0,156
Realizó tareas del hogar junto con el niño o niña	0,360	0,463	-0,103
Usó momentos de la vida cotidiana para estimular al niño o niña (ejemplo: cantarle durante el baño, masaje durante la muda)	0,640	0,741	-0,101
Apoyo social			

	Tratamiento	Control	Diferencia
Confío en mis pares (otras madres, padres o cuidadores) y puedo aprender de ellos	0,720	0,667	0,053
Tengo a alguien en quien confío para pedirle consejos sobre cómo ser madre, padre o cuidador	0,800	0,685	0,115
Tengo alguien cercano a quien pedir apoyo en momentos donde me he sentido estresado/a o sobrepasado/a con la crianza	0,800	0,722	0,078
En las últimas dos semanas me he sentido solo/a o aislado/a	0,280	0,241	0,039
Habilidades de buen trato			
Logro comprender las necesidades de mi niño o niña	0,960	0,889	0,071
Tengo herramientas para manejar situaciones desafiantes en la crianza	0,800	0,852	-0,052
Respondo de manera paciente y con un buen trato cuando mi niño o niña se porta mal o se equivoca	0,800	0,870	-0,070
Logro comprender las emociones de mi niño o niña y puedo ayudarlo/a a manejarlas de buena manera	0,880	0,852	0,028
Logro comprender que hay detrás de la conducta de mi niño o niña o lo que busca expresar con esa conducta	0,800	0,722	0,078
Dificultades parentales			
Atender a mi hijo/a a veces me quita más tiempo y energía de la que tengo	0,560	0,389	0,171
Me resulta difícil equilibrar diferentes responsabilidades debido a mi hijo/a	0,440	0,241	0,199*
El comportamiento de mi hijo/a a menudo me resulta incómodo o estresante	0,160	0,167	-0,007
Me siento abrumado/a por la responsabilidad de ser padre/madre	0,120	0,093	0,027
La mayor fuente de estrés en mi vida es mi hijo/a	0,120	0,000	0,120*
¿Cree que su niño o niña tiene un desarrollo igual al de otros niños o niñas de su edad?	0,640	0,796	-0,156
Observaciones	25	54	

Fuente: elaboración propia. Nota: Los asteriscos indican el nivel de significancia estadística de las diferencias observadas: * $p<0.1$ ** $p<0.05$ *** $p<0.001$

5.2.2 Cumplimiento (grupo que termina el proceso de evaluación)

En cuanto a la participación de las díadas de la submuestra que culmina su proceso de investigación, en la Tabla 16 se muestra su asistencia en las sesiones del programa. A diferencia de lo observado con la muestra completa, donde se evidenciaba que el 70% asistía a al menos una sesión, en este caso este porcentaje alcanza el 90%. Adicionalmente se observa que el 84% asiste a 5 sesiones o más, dando cuenta de una participación superior a la observada en la muestra completa.

Tabla 16. Participación en tratamiento (grupo que termina el proceso de evaluación)

Número de asistencia	N	%	Acumulado
No asiste	2	8,0%	
2	1	4,0%	4,0%
4	1	4,0%	8,0%
5	3	12,0%	20,0%
6	1	4,0%	24,0%
7	4	16,0%	40,0%

Número de asistencia	N	%	Acumulado
8	9	36,0%	76,0%
9	2	8,0%	84,0%
10	2	8,0%	92,0%
Total	25	100,0%	100,0%

Fuente: elaboración propia.

Por su parte, en la Tabla 17 se observa un leve aumento en la asistencia de la comuna de Quinta Normal con respecto a lo reportado en la Tabla 11 (de 2,5 a 3,7). Sin embargo, se sigue observando una importante diferencia en relación al resto de las comunas. En cuanto al promedio de la totalidad de comunas, se observa un promedio de asistencia de 7,2 sesiones y una mediana de 8 asistencias.

Tabla 17. Estadística descriptiva de asistencia (grupo que termina el proceso de evaluación)

	Concepción	El Bosque	Quinta Normal	Total
Promedio de sesiones asistidas	7,8	7,6	3,7	7,2
Mínimo	5	5	2	2
Máximo	10	10	5	10
Mediana	8,0	8,0	4,0	8,0
Cantidad de observaciones	13	7	3	23

Fuente: elaboración propia.

Por último, en la Tabla 18 se observa que el 76% de aquellas diadas que completaron el proceso de investigación, descargó por lo menos un material en la totalidad de veces que fueron enviados. En cuanto a aquellas diadas que durante todo el proceso no descargaron ningún material, alcanzan apenas un 4% de la muestra.

Tabla 18. Descarga de material (grupo que termina el proceso de evaluación)

Cantidad de veces que descarga al menos un material	N	%
No descarga ningún material	1	4,0%
2	1	4,0%
7	2	8,0%
9	2	8,0%
10	19	76,0%
Total general	25	100,0%

Fuente: elaboración propia.

Si bien los datos analizados son aún insuficientes para sacar conclusiones, los resultados preliminares de esta sección, muestran indicios de que, en general, aquel grupo que termina su proceso de investigación, es un grupo que tiene niveles de participación más altos que aquel grupo que, por diferentes razones, abandona la investigación. Esto levanta la alerta de que la atrición no es aleatoria, sino que esta relacionada con ciertas características propias de aquellas diadas que abandonan la investigación. En el caso de que estas características sean iguales para los grupos tratamiento y control, esta atrición no debiese generar problemas de validez interna. De todas formas, los análisis

de robustez planteados en la sección 4 del capítulo III tienen como objetivo descartar posibles sesgos generados por la atrición.

4. Conclusiones y desafíos del estado de avance

El presente informe da cuenta del estado de avance de la evaluación experimental del programa Crecer Jugando. En este análisis preliminar se proporciona una serie de hallazgos importantes de relevar:

1. **Resguardo de la validez interna de la evaluación:** La asignación aleatoria de las diadas (niño/a y cuidador/a) al grupo de tratamiento y al grupo de control se ha podido realizar de manera exitosa, de manera de resguardar que las diferencias observadas en los resultados pueden ser atribuidas al programa y no a factores externos, fortaleciendo la validez interna de la investigación,
2. **Desafíos de la Atrición:** Uno de los principales desafíos encontrados ha sido la atrición de participantes, tanto en la Etapa 1 como en la Etapa 2. Esta atrición, especialmente en el grupo tratamiento durante la aplicación de los instrumentos finales, plantea un riesgo de sesgo en los resultados y en el logro del número de observaciones proyectado en el protocolo inicial. Las diferencias observadas en los niveles de atrición entre comunas y entre grupos de tratamiento y control están siendo analizadas y abordadas a partir de diferentes acciones, lo que refleja una respuesta proactiva por parte del equipo de Infancia Primero y Focus ante este desafío.
3. **Resultados de Balance:** Los análisis de balance entre los grupos de tratamiento y control han mostrado que, en su mayoría, las características de las diadas son comparables, lo que respalda la equidad del proceso de aleatorización. Sin embargo, algunas diferencias significativas en variables específicas, como la frecuencia de interacción y el género de los niños/as, deben ser monitoreadas y controladas en el análisis final para evitar sesgos en la estimación del efecto del programa.
4. **Efecto Mínimo Detectable:** El cálculo del MDES ha demostrado que, a pesar de la atrición, es posible detectar efectos del programa a nivel global, aunque la sensibilidad para identificar efectos en subgrupos específicos podría ser limitada. Este hallazgo subraya la importancia de agotar todas las posibilidades para maximizar la tasa de participación para asegurar un tamaño de muestra adecuado y minimizar la atrición para obtener resultados robustos y generalizables.

En resumen, abordar los desafíos de la atrición y el sesgo de selección, junto con la mejora continua de las estrategias de implementación y análisis, serán fundamentales para maximizar la eficacia de la evaluación de Impacto del Programa Crecer Jugando. El monitoreo de estos resultados y avances serán continuamente reportados en las reuniones semanales con el equipo de Infancia Primero y en las reuniones mensuales tripartitas con el equipo de SUMMA.

V. Bibliografía

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*.

Duflo, E., Glennerster, R., & Kremer, M. (2007). *Using randomization in development economics research: A toolkit*. *Handbook of development economics*

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.

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.

Iválua (2009) *Guía práctica 4: Evaluación de la implementación*. Instituto Catalán de Evaluación de Políticas Públicas

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. doi: 10.1037/0003-066X.61.2.99. PMID: 16478355.

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.