

**MANUAL DE USUARIO
BASE DE DATOS**

**SEGUNDA RONDA
ENCUESTA LONGITUDINAL
DE LA PRIMERA INFANCIA**

Departamento de Economía
Universidad de Chile

Diagonal Paraguay 257, torre 26,
of. 1601 · Santiago, Chile

☎ (56-2) 978 34 21

www.microdatos.cl

INDICE

I. Introducción.....	3
II. Descripción General de las Bases de Datos	4
II.1 Base Hogar.....	7
II.2 Base Entrevistado	9
II.3 Base Cuidado Infantil	9
II.4 Base Historia Laboral	10
II.5 Base Evaluaciones.....	11
II.4.1 Variables de Instrumentos: Bases de Datos Evaluaciones	17
II.4.2 Edad del(de la) niño(a) seleccionado(a).....	18
III. Vinculación de Bases de Datos	19
III.1 Vinculación entre las Bases de Datos: Cohorte Transversal.....	19
III.2 Vinculación entre las Bases de Datos: Longitudinal	20
IV. Factores de Expansión.....	26
V. Referencias	28

I. Introducción

La Encuesta Longitudinal de la Primera Infancia (ELPI) es un esfuerzo conjunto entre el Gobierno de Chile y el Centro de Microdatos de la Universidad de Chile para producir información estadística que permita caracterizar el desarrollo infantil en Chile. Como tal, esta encuesta longitudinal, recoge información de los(las) niños(as) seleccionados(as) y sus hogares en diferentes momentos del tiempo, con el objetivo de caracterizar la dinámica del desarrollo infantil.

En las secciones siguientes se explican aspectos relevantes sobre el uso, contenido y la adecuada utilización de las bases de datos.

II. Descripción General de las Bases de Datos

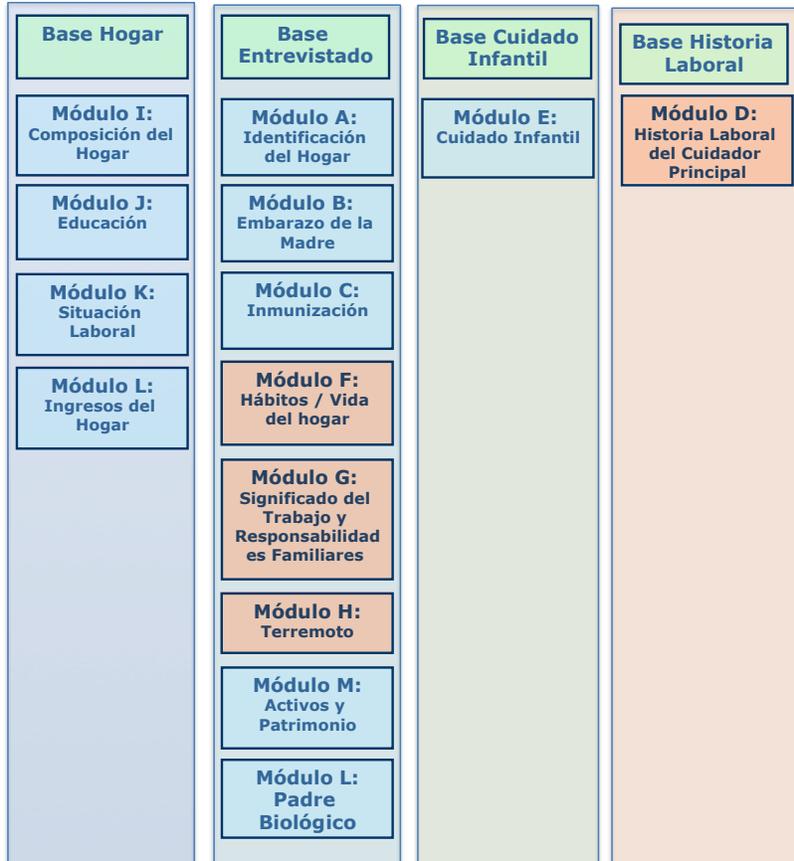
La primera ronda del estudio fue realizada el 2010 y sus resultados fueron resumidos en 4 bases de datos distintas entre sí. En ellas está contenida la información recogida a partir del **cuestionario** aplicado a los hogares con información socioeconómica del hogar del(de la) niño(a) seleccionado(a), la caracterización del cuidado infantil del mismo durante sus primeros años de vida, la caracterización de temas relevantes a partir de diferentes preguntas dirigidas al(a la) cuidador(a) principal del(de la) niño(a) y una base de datos que contiene la batería de **instrumentos** realizadas a los(las) niños(as) y su cuidador(a) principal.

En la segunda ronda del estudio, se decidió agregar un nuevo módulo relativo a la historia laboral del(de la) cuidador(a) principal. Esta información se entrega en una quinta base de datos. Esta nueva base contiene la historia laboral de cada cuidador(a) principal desde Enero de 2004 (la información se declara de manera retrospectiva).

Junto con lo anterior, en la segunda ronda de la encuesta se decidió ampliar la muestra de niños(as) para incluir a aquellos(as) nacidos(as) entre el 1º de Septiembre de 2009 y el 31 de Diciembre de 2011. Así, a la muestra longitudinal de 15.000 niños(as) se le incorporaron 3.000 niños(as) aproximadamente, los cuales participaron por primera vez en la encuesta. Por lo tanto, existen dos tipos de muestra: muestra panel y muestra nueva.

El cuestionario de la segunda ronda de la ELPI está compuesto por 14 módulos según muestra la Figura II.1. Éstos a su vez están organizados en 4 bases de datos.

Figura II.1: Esquema de Módulos Cuestionario y Bases de Datos 2012



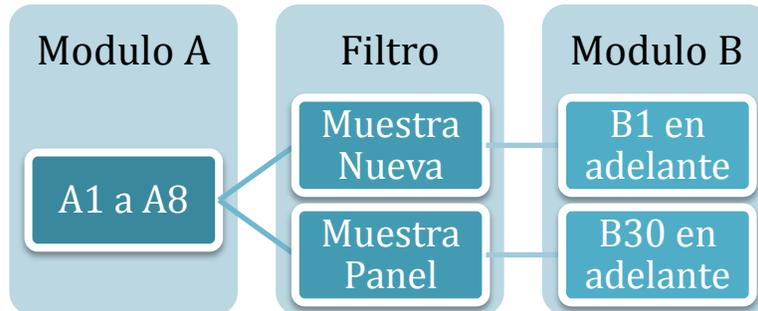
Aquellos módulos que se encuentran con color anaranjado corresponden a los que fueron incorporados en relación al cuestionario 2010. Por su parte, la información de la batería de instrumentos aplicados a niños(as) y cuidadores(as) se consolida en una sola base de datos.

Reestructuración de los módulos del cuestionario

El módulo relativo al embarazo de la madre se ubica al principio del cuestionario ya que los entrevistados eran esencialmente de dos tipos, aquellos que pertenecen a la muestra panel o de seguimiento y los entrevistados nuevos. Esta característica de la muestra y la intención de no volver a realizar un conjunto de preguntas que ya habían sido preguntadas en la primera ronda explican porque el módulo B fue ubicado

inmediatamente después del módulo A “Identificación del Hogar”. En particular la Figura II.2 detalla el flujo que siguió el encuestador al momento de realizar la entrevista.

Figura II.2: Flujo de Preguntas Módulo B para entrevistadas Panel y Nuevas



A diferencia de la encuesta del 2010, en la segunda ronda, los módulos que son contestados por todos los integrantes del hogar (Módulos: I, J, K, L y M) se ubican al final del cuestionario.

La información total de la segunda ronda de la ELPI ha sido organizada en 5 archivos diferentes:

- Cuestionario:
 - Base Hogar: contiene la información socioeconómica del hogar del(de la) niño(a) seleccionado(a);
 - Base Entrevistado: contiene la caracterización de temas relevantes a partir de diferentes preguntas dirigidas al(a la) cuidador(a) principal del(de la) niño(a);
 - Base Cuidado Infantil: contiene la caracterización del cuidado infantil del(de la) niño(a) durante sus primeros años de vida;
 - Base Historia Laboral: contiene la información de cada historia laboral del(de la) cuidador(a) principal del(de la) niño(a) desde Enero 2004 a la fecha de la entrevista.
- Instrumentos:
 - Base Evaluaciones: contiene la batería de instrumentos realizados a los(las) niños(as) y su cuidador(a) principal.

Las bases de datos contienen una variable llamada muestra que identifica a los hogares que pertenecen a la muestra panel o a la muestra nueva. La muestra panel contiene 12.898 encuestas aplicadas a los(las) cuidadores(as) principales de los(las) niños(as) seleccionados utilizando el cuestionario sociodemográfico y 11.692 evaluaciones realizadas tanto a los(las) niños(as) seleccionados(as) como a sus cuidadores(as) principales. Además, las bases de datos de la muestra de niños(as) que participan por primera vez del estudio (muestra nueva) contiene 3.135 entrevistas y 2.746 evaluaciones.

Los archivos se encuentran en formato STATA y SPSS, y contienen la siguiente información:

Cuadro II.1: Bases de Datos ELPI 2012

Nombre	Tamaño	Registros	Variables	Contenido
Hogar	11,0 MB	76.977	87	Módulo I, J, K y L
Entrevistada	9,9 MB	16.033	568	Módulo A, B, C, Módulo D (preguntas D17 a D43), Módulo E (preguntas E31 a E34), Módulo F, G, H M, N
Historia Laboral	2,1 MB	44.862	25	Módulo D (preguntas D1 a D16)
Cuidado Infantil	3,9 MB	60.605	52	Módulo E (preguntas E1 a E30)
Evaluaciones	19,8 MB	14.438	203	Todas las evaluaciones

A continuación se describe cada una de las bases de datos.

II.1 Base Hogar

Esta base de datos contiene las preguntas de los módulos I, J, K y L. Cada una de las preguntas de estas secciones se realizan al(a la) cuidador(a) principal del(de la) niño(a) seleccionado(a), el(la) que a su vez responde cada pregunta para cada uno de los miembros del hogar. Por lo tanto, en esta base de datos cada fila representa un miembro del hogar.

La base contiene la variable **folio** que identifica a qué hogar pertenece el(la) niño(a) seleccionado(a) y contiene una variable llamada **orden** que identifica a cada miembro al

interior del hogar. Adicionalmente existe una variable llamada **id_persona** que combina el folio y el orden generando un identificador único por persona. Por último y para facilitar el uso de esta base, se agregó una variable **tot_per** que asigna, para todos los miembros del hogar, el número de personas que viven en ese hogar.

Por ejemplo, la siguiente tabla muestra la lógica de la base de datos hogar, el folio identifica al hogar, por lo tanto todas las filas donde el folio se repite corresponden al mismo hogar. Luego la variable orden es un sub-código al interior de cada folio que identifica a cada miembro del hogar, donde orden 1 siempre es el(la) cuidador(a) principal. Entonces, en el ejemplo el hogar con folio 100011 está conformado por cuatro personas (incluido el mismo cuidador(a)), y el hogar del(de la) cuidador(a) con folio 100041 está formado por dos personas, el(la) cuidador(a) y el(la) niño(a).

Cuadro II.2: Ejemplo Bases de Datos Hogar 2012

folio	orden	id_persona	tot_per
100011	1	1000101	4
100011	2	1000102	4
100011	3	1000103	4
100011	4	1000104	4
100041	1	10004101	2
100041	2	10004102	2

Las secciones del cuestionario que están contenidas en estos módulos son aquellas que recogen información relacionada con características demográficas y educacionales de los miembros del hogar, la situación laboral y de ingresos de cada uno de ellos¹.

Las preguntas de la base hogar, de los módulos I y J se realizan para todos los miembros del hogar, mientras que las preguntas de los módulos K y L se realizan sólo para los integrantes del hogar que son mayores de 15 años.

Las preguntas en la base de datos se identifican con la letra de cada módulo y luego con

¹ Para mayor descripción de los temas y preguntas incorporadas en estos módulos ir al informe "Presentación Segunda Ronda Encuesta Longitudinal de la Primera Infancia" y el "Libro de Códigos 2012", ambos disponibles en la página web www.elpi.cl.

el número de cada pregunta del cuestionario.

II.2 Base Entrevistado

Esta base de datos contiene las preguntas de los módulos A, B, C, F, G, H, M y N, además de las preguntas D17 a D43 del módulo D y preguntas E31 a E34 del módulo E². Cada una de las preguntas de estas secciones se realizan al(a la) cuidador(a) principal del(de la) niño(a) seleccionado(a), por lo tanto, en esta base de datos cada fila representa un individuo, los que se pueden identificar mediante la variable folio.

Las preguntas de la base entrevistado, se realizan para todos los(las) cuidadores(as), excepto en el módulo B, donde las preguntas B1 a B18 sólo se realizan para aquellas cuidadoras principales de la muestra nueva que son madres biológicas del(de la) niño(a) seleccionado(a). Desde la pregunta B19 a B29 las preguntas están dirigidas para todos(as) los(las) cuidadores(as) principales de la muestra nueva y desde la B30 en adelante contestan todos(as) los(las) cuidadores(as), ya sean muestra nueva o muestra panel. En el caso del módulo N, es respondido solo en los casos en que el padre biológico del(de la) niño(a) seleccionado(a) no vive en el hogar.

Las preguntas en la base de datos se identifican con la letra de cada módulo y luego con el número de cada pregunta del cuestionario.

Adicionalmente en esta base se agregan la variables de región y área, indicando en que región y zona (urbano/rural) se encuentra el hogar del(de la) niño(a).

II.3 Base Cuidado Infantil

Esta base contiene las preguntas del Módulo E Cuidado Infantil. Estas preguntas son las que el(la) cuidador(a) principal debe responder para cada uno de los tramos de edad del(de la) niño(a) hasta completar el tramo correspondiente a la edad del(de la) niño(a) en el momento de la entrevista.

² Para mayor descripción de los temas y preguntas incorporadas en estos módulos ir al informe "Presentación Segunda Ronda Encuesta Longitudinal de la Primera Infancia" y el "Libro de Códigos 2012", ambos disponibles en la página web www.elpi.cl.

Los tramos de edad son entre 0 - 3 meses, 3 - 6 meses, 6 - 12 meses, 12 - 18 meses, 2 - 3 años, 3 - 4 años, 4 - 5 años, 5 - 6 años y entre 6 - 7 años. Cada cuidador(a) debe reportar según el tramo etario del(de la) niño(a) hasta el tramo correspondiente a la edad actual. Por lo tanto, hay más de una fila por entrevistado y cada fila representa un tramo etario. La variable **folio** es la que identifica al(a la) niño(a) y la variable **orden** es la que identifica cada tramo de edad.

Este módulo se preguntó por primera vez el 2010 a todos(as) los(las) niños(as) de la muestra panel. Por lo tanto, en la segunda ronda, se preguntó nuevamente el módulo pero solamente aquellos tramos de edad en que faltaba información. Por ejemplo, en el siguiente cuadro, el(la) entrevistado(a) con folio 100011 corresponde a un(a) niño(a) de la muestra nueva y el entrevistado 500011 corresponde a un niño(a) de la muestra panel. El(La) cuidador(a) del(de la) niño(a) de la muestra panel ya respondió para los tramos anteriores a 4-5 años en la primera ronda, por lo tanto ahora solo se le pregunta por aquellos tramos de edad que han transcurrido entre el 2010 y el 2012. Por otro lado, al(a la) niño(a) de la muestra nueva (folio 500011) se le pregunta por cada uno de los tramos hasta su edad actual.

Cuadro II.3: Bases de Datos Cuidado Infantil 2012

Muestra	Folio	Orden	Tramo	Pregunta 1
Nueva	500011	1	0 - 3 meses	3
Nueva	500011	2	3 - 6 meses	2
Nueva	500011	3	6 - 12 años	3
Panel	100011	1	4 - 5 años	2
Panel	100011	2	5 - 6 años	.

II.4 Base Historia Laboral

Esta base contiene la mayoría de las preguntas del Módulo D (Historia Laboral del(de la) Cuidador(a) Principal). En estas preguntas el(la) cuidador(a) principal debe responder respecto de su situación laboral en determinados momentos del tiempo³. Dependiendo de cuál sea la situación del(de la) cuidador(a) principal: trabajando, cesante, buscando

³ En este módulo se pregunta desde Enero de 2004 hasta el mes en que se realizó la entrevista.

trabajo por primera vez, inactivo(a) o menor de 15 años, continuará o no con las preguntas del módulo. Por ejemplo, en el cuadro siguiente el(la) cuidador(a) principal se encontraba trabajando desde Enero de 2004 hasta Junio de 2008, momento en el cual el(ella) declara que dejó de trabajar y pasó a estar cesante hasta Diciembre de 2011. Finalmente, en Enero de 2012 volvió a encontrar trabajo hasta el momento en que se realizó la entrevista en el mes de Julio de 2012.

Cuadro II.4: Bases de Datos Historia Laboral 2012

Folio	Orden	d1im	d1ia	d1tm	d1im	D2
100011	1	01	2004	06	2008	1
100011	2	07	2008	12	2011	2
100011	3	01	2012	07	2012	1

En las cuatro bases descritas anteriormente, existen variables que tienen códigos de respuesta no incorporados en el cuestionario pero que debido a las respuestas de los entrevistados debieron agregarse en la base de datos. Para mayor detalle de estos códigos ir a Libro de Códigos 2012 disponible en la página web www.elpi.cl.

Por otra parte, en algunos casos fue necesario realizar un proceso de codificación para preguntas abiertas o aquellas en que permitían alternativa Especifique. Estas variables tienen el mismo nombre que la original pero con el sufijo "cod".

II.5 Base Evaluaciones

Esta base de datos contiene las preguntas de los distintos instrumentos aplicados tanto al(a la) cuidador(a) principal como al(a la) niño(a) seleccionado(a), abarcando las áreas desarrollo psicomotor, función ejecutiva (en revisión - pendientes), socioemocional y medidas antropométricas para los(as) niños(as) y áreas cognitiva, socioemocional y medidas antropométricas para el(la) cuidador(a) principal. El Cuadro II.5 resume los instrumentos aplicados al(a la) niño(a) según rango de edad, mientras que el Cuadro II.6 resume los instrumentos aplicados al(a la) cuidador(a) principal.

Cuadro II.5: Instrumentos Aplicados al(a) Niño(a) por Rango de Edad

Área	Instrumento	Tramo de edad
Desarrollo Psicomotor	BDI-2 (Inventario de Desarrollo Battelle – Screening Test)	6 – 83 meses, 30 días
	TADI (Test de Aprendizaje de Desarrollo Infantil)	6 – 83 meses, 30 días
	TVIP (Test de Vocabulario en Imágenes Peabody)	30 – 83 meses, 30 días
Función Ejecutiva	SDT (Snack Delay Task)	24 - 35 meses, 30 días (2 - 3 años)
	PTT (Pencil Tapping Task)	24 - 35 meses, 30 días (2 - 3 años)
	BDS (Backward Digit Span Task)	36 – 83 meses, 30 días
	HTKS (Head Toes Knees Shoulders Task)	36 – 83 meses, 30 días
Socioemocional	ASQ:SE. 6 meses	6 – 8 meses, 30 días
	ASQ:SE. 12 meses	9 – 14 meses, 30 días
	ASQ:SE. 18 meses	15 –17 meses, 30 días
	CBCL1	18- 71 meses, 30 días
	CBCL2	72- 83 meses, 30 días
Medidas Antropométricas	Altura	6 – 59 meses, 30 días
	Peso	6 – 59 meses, 30 días
	Circunferencia Craneal	6 – 59 meses, 30 días

Cuadro II.6: Instrumentos Aplicados a la Muestra por Rango de Edad

Área	Instrumento	Tramo de edad
Cognitiva	Escala Dígitos WAIS	Muestra nueva y menores de 20 años
	Escala Vocabulario WAIS	Muestra nueva y menores de 20 años
Socioemocional	<i>Big Five Inventory (BFI)</i>	Muestra nueva y menores de 20 años
	Escala de Depresión Post Parto Edimburgo	Madres de niños entre 6 meses y 1 año.
	<i>Parent Stress Index (PSI)</i>	Para toda la muestra
Medidas Antropométricas	Altura	Para toda la muestra
	Peso	Para toda la muestra

A continuación se enumeran los instrumentos evaluados en la segunda ronda de la ELPI junto con una explicación de cómo están presentadas las variables de los mismos en la base de datos. Para una descripción detallada de los instrumentos ver informe "Presentación Segunda Ronda Encuesta Longitudinal de la Primera Infancia", para una presentación de los principales resultados de los instrumentos ver informe "Resultados Evaluaciones Segunda Ronda Encuesta Longitudinal de la Primera Infancia" y para una descripción detallada de las normalizaciones realizadas a los instrumentos ver informe "Metodológico Encuesta Longitudinal de la Primera Infancia", todos disponibles en www.elpi.cl.

A. Área Desarrollo Psicomotor del(de la) Niño(a)

El desarrollo psicomotor es el proceso por el cual le permite al(a la) niño(a) relacionarse, conocer y adaptarse al medio que lo(a) rodea. Este proceso incluye aspectos como el lenguaje expresivo y comprensivo, coordinación viso-motora, motricidad gruesa, equilibrio y el aspecto social-afectivo, que está relacionado con la autoestima. La batería

de instrumentos aplicados en la ELPI miden el desarrollo de los(las) niños(as) en las áreas motora, cognitiva, lenguaje, emocional y social. El poder evaluar el desarrollo psicomotor de los(las) niños(as) permite identificar las áreas en que éstos(as) tienen mayores y menores logros y, consecuentemente, retroalimentar los programas públicos que podrían tener un impacto en dichos indicadores de desarrollo. En la segunda ronda de la ELPI se aplicaron tres instrumentos que miden estas áreas.

- Inventario de Desarrollo Battelle – Screening Test, 2da ed. (BDI-2), (Newborg, Stock y Wnek , 1996).
- Test de Aprendizaje de Desarrollo Infantil (TADI), (Edwards et al., 2013).
- Test de Vocabulario en Imágenes Peabody, adaptación Hispanoamericana (TVIP), (Dunn, Padilla, Lugo y Dunn, 1996)

B. Área Función Ejecutiva del(de la) Niño(a)

La función ejecutiva es la función cerebral que permite establecer metas, hacer planes, iniciar tareas y operaciones mentales, y realizarlas de manera eficiente. Específicamente refiere a los procesos cognitivos de orden superior que requieren autorregulación. Su función es monitorear y controlar el pensamiento y la acción. La capacidad de regular la conducta implica la integración de las habilidades de atención, flexibilidad cognitiva, memoria de trabajo y control inhibitorio.

En múltiples estudios se ha visto que la función ejecutiva predice logros académicos (Blair y Razza, 2007; McClelland, Acock y Morrison, 2006). Debido al poder predictivo que se le atribuye a la función ejecutiva se considera que estas pruebas pueden entregar información relevante para programas que trabajan con niños(as) en primera infancia. En la segunda ronda de la ELPI se aplicaron cuatro instrumentos que miden esta área.

- Snack Delay Task (SDT), (Kochanska, Murray, Jacques, Koenig, y Vandergeest, 1996).
- Pencil Tapping Task (PTT), (Diamond y Taylor, 1995).
- Head Toes Knees Shoulders Task (HTKS), (McClelland, Cameron, Connor,

Farris, - Jewkes y Morrison, 2007).

- Backward Digit Span Task (BDS), (Davis y Pratt, 1996).

Estos resultados aun están en revisión por parte del equipo investigador por lo que solo se encuentran los nombres de las variables en las bases de datos. Una vez concluido el análisis estos serán dados a conocer.

C. Área Socioemocional del(de la) Niño(a)

El desarrollo socioemocional de los(as) niños(as) en primera infancia es fundamental para su desarrollo posterior (Siegel, 1999), teniendo un impacto no sólo en los resultados cognitivos y académicos, sino en diversos ámbitos de la vida de los individuos. En la segunda ronda de la ELPI se aplicaron dos instrumentos de evaluación que miden esta área.

- Ages & Stages Questionnaire: Social Emotional (ASQ: SE), (Squires, Bricker y Twombly, 2006).

- Child Behavior Checklist (CBCL1) y (CBCL2), (Achenbach y Rescorla, 2000)

D. Medidas Antropométricas del(de la) Niño(a)

Las medidas antropométricas contenidas en la base de datos son: el peso, la talla, el índice de masa corporal (IMC) y la circunferencia craneal (CC) tanto para los(las) niños(as) que caminan como para los que no caminan. Los(Las) niños(as) que no caminan fueron evaluados en brazos de su cuidador(a) principal. Además, la base de datos contiene el perfil o interpretación de las medidas antropométricas del(de la) niño(a) en base a los parámetros establecidos por la Organización Mundial de la Salud (OMS). Los primeros indicadores son, la interpretación del peso para la edad, la interpretación de la talla para la edad y la interpretación de la circunferencia craneal para la edad. Éstos clasifican a los(las) niños(as) en un perfil de "retraso", "normal" o "aumentado", según los estándares de la OMS. Los segundos indicadores corresponden al perfil o interpretación del IMC y la interpretación de la talla para el peso. El primero clasifica a los(las) niños(as) según su estado nutricional: "bajo peso", "normal" o "sobre peso". El segundo en cambio, clasifica a los(las) niños(as) en un perfil de "retraso", "normal" o "aumentado", dependiendo de su desarrollo en esta dimensión.

E. Área Cognitiva del(de la) Cuidador(a) Principal

En este estudio se seleccionaron dos sub escalas representativas del instrumento Wechsler Adults Intelligence Scale (WAIS): Sub escala Retención de Dígitos y Vocabulario (Apfelbeck y Hermosilla, 2000).

F. Área Socioemocional del(de la) Cuidador(a) Principal

El instrumento evaluado en el area socioemocional fue el *Big Five Inventory* (BFI), (Casullo, 2000; Tupes y Christal, 1961), el cual presenta resultados para las siguientes 5 dimensiones: Neuroticismo, Extraversión, Apertura a la experiencia, Amabilidad, Responsabilidad. Además, se los instrumentos Escala de Depresión Postparto de Edimburgo (EPDS), (Cox, Holden y Sagovsky, 1987) y Parent Stress Index Short Form (PSI), (Abidin, 1995), fueron también aplicados con el fin de caracterizar de mejor manera esta área de desarrollo.

G. Medidas Antropométricas del(de la) Cuidador(a) Principal

En la base de datos fueron incluidas las siguientes medidas antropométricas para los(las) cuidadores(as) principales del(de la) niño(a) seleccionado(a): peso, talla, índice de masa corporal (IMC) y estado nutricional. El estado nutricional del(de la) cuidador(a) principal clasifica a la persona en "bajo peso", "normal", "sobrepeso", "obesidad moderada", "obesidad severa" y "obesidad muy severa".

H. Evaluación del Hogar

En este estudio, se diseñó una escala de observación basada en el Inventario Home Observation for the Measurement of the Environment (HOME) (Caldwell y Bradley, 1984). Por lo tanto, en la base de datos están las 45 preguntas que fueron utilizadas para la aplicación de la adaptación del HOME. Éstas fueron divididas en dos grupos, de tal forma que las preguntas fueran consistentes con la edad que tenía el(la) niño(a) al momento de la entrevista. El primer grupo de preguntas son 27 y están catalogados en la base de datos con el sufijo "hm1_" , mientras que las 17 preguntas que corresponden al segundo grupo están catalogadas con el sufijo "hm2_". Es importante notar que estas

preguntas fueron etiquetadas en las base de datos de tal forma que fueran consistentes con las preguntas utilizadas en el mismo instrumento aplicado durante la primera ronda. Por ejemplo, en la base de datos la primera variable del instrumento es hm1_1, variable que tiene el mismo nombre en la base de datos de la primera ronda.

II.4.1 Variables de Instrumentos: Bases de Datos Evaluaciones

En la base de datos "Evaluaciones" se presenta el puntaje bruto, el puntaje tabla (T) o percentil cuando corresponda y la categorización o perfil que permite la interpretación de los instrumentos a partir de los resultados del mismo.

El puntaje bruto para cada uno de los instrumentos se calcula según se detalla en el informe "Presentación Primera Ronda Encuesta Longitudinal de la Primera Infancia" disponible en www.elpi.cl. Además en la base de datos todos los puntajes brutos están presentados con el sufijo "_pb", por ejemplo, **std_pb**⁴. Luego, a partir del puntaje bruto se calcula un puntaje tabla o percentil dependiendo del instrumento y de la normalización utilizada por los autores del instrumento cuando existe o por el equipo técnico del Centro de Microdatos⁵ (CMD) cuando no hay una referencia nacional o internacional. Los puntajes tabla o T están presentados con el sufijo "_pt", por ejemplo, **sdt_pt**⁶. Finalmente estos valores tabla o percentiles se trasladan a valores de corte de interpretación, los cuales son calculados según el criterio de los autores o del CMD según corresponda. Todos los puntajes de las categorías de interpretación se presentan con el sufijo "_int", por ejemplo **sdt_int**⁷. Es importante notar que en la

⁴ Notar también que cada variable en su etiqueta se explica como se define, por ejemplo, en la variable sdt_pb la etiqueta dice "SDT. Puntaje Bruto".

⁵ Para ver detalles de las normalizaciones realizadas por el equipo técnico del Centro de Microdatos ver informe " Metodológico Encuesta Longitudinal de la Primera Infancia", disponible en www.elpi.cl.

⁶ Notar que en estos casos, la etiqueta detalla como se define cada variable. Por ejemplo en la variable sdt_pt la etiqueta es "SDT. Puntaje T").

⁷ Notar que en estos casos, la etiqueta detalla como se define cada variable. Por ejemplo en la variable sdt_int la etiqueta es "SDT. Interpretación Puntaje T").

etiqueta de cada variable de interpretación se detalla en base a que normalización se realizó la categorización, es decir, tabla o percentil.

II.4.2 Edad del(de la) niño(a) seleccionado(a)

En la base de datos se agregó la edad en meses (*edad_meses*) del(de la) niño(a) seleccionado(a). Para poder calcularla se utilizó la fecha de nacimiento del(de la) niño(a), a partir de los datos del registro civil, en conjunto con la fecha en que se realizó la evaluación psicológica en el hogar del(de la) niño(a).

A continuación se detalla como fue calculada la edad en meses.

$$\begin{aligned} \textit{edad_meses} = & (2012 - \{\textit{año de nacimiento}\}_i) \times 12 \\ & + (\{\textit{mes de entrevista}\}_i - \{\textit{mes de nacimiento}\}_i) \end{aligned}$$

Debido a que la evaluación al(a la) niño(a) se realizó un tiempo después de la entrevista en el hogar, es posible que existe un desfase en la edad en meses reportada en la base evaluaciones y la edad en años reportada en la base de hogar.

III. Vinculación de Bases de Datos

III.1 Vinculación entre las Bases de Datos: Cohorte Transversal

Para que el investigador pueda utilizar la totalidad de la información es importante entender como vincular correctamente las diferentes bases de datos. Para eso es que el hogar está identificado con un número único denominado folio. El folio es único para cada encuesta por lo tanto, dependiendo de la base de datos puede representar diferentes conjuntos de datos.

- En la base hogar el folio es único para todos los integrantes del hogar.
- En la base cuidado infantil el folio es único para el conjunto de historias de cuidado infantil del(de la) niño(a).
- En la base historia laboral el folio es único para el conjunto de historias laborales del(de la) cuidador(a) principal.

En estas tres bases de datos, además existe un número correlativo dentro de cada folio cuya variable se llama orden, y que permite individualizar cada persona, historia de cuidado o historia laboral, respectivamente.

- En la base entrevistado el folio es único para cada cuidador(a).
- En la base evaluaciones el folio es único para cada niño(a) y/o cuidador(a) principal evaluado(a).

Por lo tanto, folio es la variable que permite pegar las bases de datos entre sí. En el Cuadro III.1 se presenta el esquema de vinculación entre las bases de datos de la segunda ronda de la ELPI. El objetivo de este cuadro es facilitar al investigador el manejo de la información entre las bases de datos. Para esto se definen las siguientes relaciones:

- **1:1** : En la base de datos maestra hay 1 observación por folio y en la base auxiliar hay 1 observaciones por folio.
- **1:m** : En la base de datos maestra hay 1 observación por folio y en la base auxiliar hay m observaciones por folio.
- **m:m** : En la base de datos maestra hay m observación por folio y en la base auxiliar hay m observaciones por folio.
- **m:1** : En la base de datos maestra hay m observación por folio y en la base auxiliar hay 1 observación por folio.

Cuadro III.1: Esquema de Relaciones de Vinculación entre Bases de Datos

Base de Datos Maestra	Relación	Base de Datos Auxiliar
Entrevistada	1:m	Hogar
Entrevistada	1:m	Cuidado Infantil
Entrevistada	1:m	Historia Laboral
Entrevistada	1:1	Evaluaciones
Hogar	m:m	Cuidado Infantil
Hogar	m:1	Evaluaciones
Hogar	m:m	Historia Laboral
Cuidado Infantil	m:1	Evaluaciones
Cuidado Infantil	m:m	Historia Laboral

Para ver un ejemplo de cómo construir una base de datos con un único registro por folio para una investigación sobre la primera infancia, ver “Manual de Usuario Base de Datos Primera Ronda Encuesta Longitudinal de la Primera Infancia” disponible en www.elpi.cl.

III.2 Vinculación entre las Bases de Datos: Longitudinal

Evaluaciones: La llave que permite vincular las evaluaciones realizadas el 2010 con las evaluaciones realizadas el 2012 es el folio.

Además es importante mencionar que debido a que la evaluación se realizaba días después de la entrevista en el hogar, era factible que (el)la cuidador(a) principal estuviera presente en la entrevista y no estuviera presente en la evaluación del(de la)

niño(a). Es por esto que en esta base se incluyó un identificador llamado *folioev*, el cual vincula 1:1 la persona que estuvo presente en la evaluación del(de la) niño(a) con los integrantes del hogar contenidos en la Base Hogar. Esta vinculación se debe realizar a través de la variable *folioev* (Base Evaluaciones) y la variable *id_persona* (Base Hogar).

Encuestas: La variable que permite vincular 1:1 las personas de la base hogar entre la primera ronda y la segunda ronda de la ELPI, debe ser construida de la siguiente forma:

En la Base Hogar 2010: $\text{foliopersona} = \text{folio} * 100 + \text{orden}$

En la Base Hogar 2012: $\text{foliopersona} = \text{folio} * 100 + i_orden$

La variable *i_orden* contiene el valor numérico de la variable *orden* de la primera ronda para cada uno de los integrantes del hogar.

Una vez construidas estas variables es posible utilizar la llave *foliopersona* para vincular las personas dentro del hogar en la primera ronda con las personas dentro del hogar en la segunda ronda. Naturalmente en este proceso y debido a que la composición del hogar del(de la) niño(a) seleccionado(a) cambia con el tiempo, se pueden dar los siguientes casos:

- Personas que permanecen dentro del hogar el 2010 y 2012.
- Personas que están dentro del hogar el 2010 y no están el 2012.
- Personas que no están dentro del hogar el 2010 y si están el 2012.

Por otro lado, cuando el objetivo no es pegar a las personas dentro del hogar, sino más bien generar información agregada del(de la) niño(a) entre las diferentes rondas, la llave única es la variable *folio*. Por ejemplo, si un investigador quisiera vincular el número de personas dentro del hogar del 2010 con alguna variable de su interés de la base entrevistado del 2012, como por ejemplo los hábitos de alimentación, se debe utilizar la variable *folio* como llave de identificación única.

Ejemplo: Construcción de Base de Datos Longitudinal de Niños(as) Seleccionados(as)

Supongamos que un investigador está interesado en generar una base de datos longitudinal o panel y que tenga dos observaciones por folio, es decir una observación para la primera y otra para la segunda ronda por cada niño(a) seleccionado(a). Por ejemplo, supongamos que el investigador necesita evaluar cómo se comportan estadísticamente los resultados de los instrumentos de evaluación (Base Evaluaciones) del(de la) niño(a) según variables relacionadas con el hogar (Base Hogar). Por simplicidad asumamos que al investigador le interesa saber cómo se comporta estadísticamente el resultado en el TVIP⁸, según el ingreso familiar per cápita del hogar.

A continuación se describe una forma en que el investigador podría vincular la información entre las bases de datos:

- 1) **Base Hogar 2010** (m registros por folio): Crear y agrupar toda la información necesaria del hogar, por ejemplo el ingreso per cápita del hogar, a través del folio ya que éste es único para todos los individuos del hogar. Luego, quedarse con un único registro por hogar (1 registro por folio). Guardar esta base como Base Hogar 2010 2 (ver Cuadro III.2).
- 2) **Base Hogar 2012** (m registros por folio): Crear y agrupar toda la información necesaria del hogar, por ejemplo el ingreso per cápita del hogar, a través del folio ya que éste es único para todos los individuos del hogar. Luego, quedarse con un único registro por hogar (1 registro por folio). Guardar esta base como Base Hogar 2012 2 (ver Cuadro III.4).
- 3) **Base Evaluaciones 2010** (1 registro por folio): Generar una base que contenga solo los casos validos del instrumento TVIP para los(las) niños(as) evaluados el 2010. Es importante recordar que no todos los(las) niños(as) que participaron en la encuesta fueron evaluados. Guardar esta base como Base Evaluaciones 2010 2 (ver Cuadro III.3).
- 4) **Base Evaluaciones 2012** (1 registro por folio): Generar una base que contenga

⁸ Para una completa descripción de este instrumento ver informe "Metodológico Encuesta Longitudinal de la Primera Infancia" disponible en la página web www.elpi.cl.

solo los casos validos del instrumento TVIP para los(las) niños(as) evaluados el 2012. Es importante recordar que no todos los(las) niños(as) que participaron en la encuesta fueron evaluados. Guardar esta base como Base Evaluaciones 2012 2 (ver Cuadro III.5).

Cuadro III.2: Base de Datos
Hogar 2 2010 (1:1)

Folio	Ingreso Per Cápita 2010
100011	50.000
100051	150.000
100081	450.000
100121	27.000

Cuadro III.3: Base de Datos
Evaluaciones 2 2010 (1:1)

Folio	TVIP 2010
100011	26
100081	30
100121	80

Cuadro III.4: Base de Datos
Hogar 2 2012 (1:1)

Folio	Ingreso Per Cápita 2012
100011	70.000
100081	320.000
100121	60.000

Cuadro III.5: Base de Datos
Evaluaciones 2 2012 (1:1)

Folio	TVIP 2012
100011	45
100081	50
100121	75

Nota: El folio 100051 si fue encuestado el 2010, pero no fue encuestado el 2012, por lo tanto dicho niño(a) no está contenido en las bases de datos 2012.

- 5) **Base Ronda 2010** (1 registro por folio): A través de los programas estadísticos tradicionales (SAS, Stata, SPSS, entre otros), vincular la "Base Hogar 2010 2" con la "Base Evaluaciones 2010 2". Generar, a partir de esta vinculación, una base "Ronda 2010" (ver Cuadro III.6)

Cuadro III.6: Base de Datos Ronda 2010 (1:1)

Folio	TVIP 2010	Ingreso Per Cápita 2010
100011	26	50.000
100051	.	150.000
100081	30	450.000
100121	80	27.000

Nota: El folio 100051, no tiene dato en el TVIP 2010, porque pese a que se realizó la encuesta en el hogar no se realizó la evaluación.

- 6) **Base Ronda 2012** (1 registro por folio): A través de los programas estadísticos tradicionales (SAS, Stata, SPSS, entre otros), vincular la "Base Hogar 2012 2" con la "Base Evaluaciones 2012 2". Generar, a partir de esta vinculación, una base "Ronda 2012" (ver Cuadro III.7).

Cuadro III.7: Base de Datos Ronda 2012 (1:1)

Folio	TVIP 2012	Ingreso Per Cápita 2012
100011	45	70.000
100081	50	320.000
100121	75	60.000

- 7) **Base Longitudinal Versión 1** (1 registro por folio): A través de los programas estadísticos tradicionales (SAS, Stata, SPSS, entre otros), vincular mediante el folio la "Base Ronda 2010" con la "Base Ronda 2010".

Cuadro III.8: Base de Datos Longitudinal Versión 1 (1:1)

Folio	TVIP 2010	Ingreso Per Cápita 2010	TVIP 2012	Ingreso Per Cápita 2012
100011	26	50.000	45	70.000
100051	.	150.000	.	.
100081	30	450.000	50	320.000
100121	80	27.000	75	60.000

- 8) **Base Longitudinal Versión 2** (2 registros por folio): Finalmente y a través de los programas estadísticos tradicionales (SAS, Stata, SPSS, entre otros), reestructurar la "Base de Datos Longitudinal Versión 1" de tal forma que queden

dos observaciones por folio, donde la primera observación corresponde a los datos de la primera ronda y la segunda observación corresponde a los datos de la segunda ronda (ver Cuadro III.9).

Cuadro III.9: Base de Datos Longitudinal Versión 2 (2:1)

Folio	Año	TVIP	Ingreso Per Cápita
100011	2010	26	50.000
100011	2012	45	70.000
100051	2010	.	150.000
100051	2012	.	.
100081	2010	30	450.000
100081	2012	50	320.000
100121	2010	80	27.000
100121	2012	75	60.000

A través de estos pasos el investigador contará con un base de dos registros por folio que contiene la información relacionada al instrumento de evaluación TVIP y las características del hogar del(de la) niño(a), en particular el ingreso familiar per cápita. Es importante notar que existen en este caso dos tipos de no respuestas a la unidad, que deben ser consideradas, atrición y no respuesta al instrumento de evaluación. La atrición, se conoce como la no respuesta asociada a la encuesta entre diferentes rondas del estudio. Mientras que la no respuesta al instrumento ocurre porque un porcentaje de los(las) niños(as) no fue evaluado en alguna ronda, pese a realizarse la encuesta.

IV. Factores de Expansión

A todas las bases se le agregaron los respectivos ponderadores, que dan cuenta del diseño muestral y las correcciones efectuadas por no respuesta⁹. Debido a que el estudio está compuesto por dos tipos de muestras, niños(as) que participan por primera y por segunda vez, es necesario la utilización de dos tipos de factores de expansión: factores de expansión de cohorte transversal y factores de expansión longitudinales o de panel.

Factores de Expansión de Cohorte Transversal

- fexp_enc0: Factor de expansión de cohorte transversal para las entrevistas.
- fexp_test0: Factor de expansión de cohorte transversal para las evaluaciones.
- fexp_hog0: Factor de expansión de cohorte transversal para los hogares.

Factores de Expansión Longitudinal

- fexp_encP: Factor longitudinal para las entrevistas.
- fexp_hogP: Factor longitudinal para los hogares.
- fexp_testP: Factor longitudinal para las evaluaciones.

El factor de expansión que se debe utilizar para expandir a nivel nacional la muestra de hogares es **fexp_hog0**, éste permite estimar el total de personas en hogares con niños y niñas nacidos entre el 1º de Enero de 2006 y el 31 de Diciembre de 2011. Por otro lado, si el investigador está interesado en utilizar el carácter longitudinal de la base hogar, debe hacer uso del factor **fexp_hogP**.

Debido a que el número de evaluaciones obtenidas es menor al número de entrevistas realizadas, existe un factor para hacer análisis utilizando tanto la información de la base entrevistadas como de la base evaluaciones. Para utilizar la totalidad de evaluaciones realizadas en la segunda ronda es decir las 11.692 evaluaciones correspondientes a la

⁹ Para mayor detalle de los factores de expansión ir a informe "Diseño Muestral y Factores de Expansión Segunda Ronda Encuesta Longitudinal de la Primera Infancia" disponibles en www.elpi.cl.

muestra panel más las 2.746 evaluaciones de la muestra nueva se debe utilizar el factor de cohorte transversal **fexp_test0**, mientras que para las evaluaciones realizadas en la muestra longitudinal se debe utilizar el factor **fexp_testP**. Este último factor de expansión pondera las 11.198 evaluaciones realizadas en ambas rondas para obtener una estimación poblacional de la totalidad del marco muestral.

V. Referencias

- Abidin, R.R. (1995). "Parenting Stress Index (PSI) manual (3rd ed.). Charlottesville, VA: Pediatric Psychology Press.
- Achenbach, T. M., & Rescorla, L. A. (2000). "Manual for the ASEBA Preschool Forms & Profiles". Burlington: University of Vermont, Research Center for Children, Youth & Families.
- Apfelbeck, E.-M., y Hermosilla, M. (2000). "Manual de Administración y Tabulación del Test de WAIS". Santiago.
- Blair C, Razza RP. (2007). "Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten". *Child Dev.* Mar-Apr;78(2):647-663. PubMed PMID: 17381795.
- Caldwell, B. y Bradley, R. (1984). "Home Observation for Measurement of the Environment (HOME) - Revised Edition". University of Arkansas, Little Rock.
- Casullo, M. (2000). "Big Five Inventory". Buenos Aires, Argentina: Universidad de Buenos Aires.
- Cox, J.L., Holden, J.M., and Sagovsky, R. (1987). "Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale". *British Journal of Psychiatry* 150:782-786.
- Davis, H. L., & Pratt, C. (1996). "The development of children's theory of mind: The working memory explanation". *Australian Journal of Psychology*, 47, 25-31.
- Diamond, A., & Taylor, C. (1996). "Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do"". *Developmental Psychobiology*, 29, 315-334.
- Dunn, L. M., Padilla, E., Lugo, D., & Dunn, L. M. (1986). *Manual del examinador para el Test de Vocabulario en Imágenes Peabody*. Minneapolis: Pearson.

- Edwards, M.; Gomez, M. Pardo, M. (2012). "Test de Aprendizaje y Desarrollo Infantil". Working document no publicado. CEDEP
- McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, M., & Morrison, F. J. (2007). "Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills". *Developmental Psychology*, 43, 947–959
- McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). "The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school". *Early Childhood Research Quarterly*, 21(4), 471 – 490
- Newbor J, Stock J, Wnek L. (1996). "Inventario de desarrollo Battelle". Madrid: TEA.
- OMS. (2010). "Patrones de crecimiento infantil"
<http://www.who.int/childgrowth/publications/>
- Siegel, D. (1999). *The developing mind*. New York: Guilford Press.
- Squires, J., Bricker, D., & Twombly, E. (2006). "The ASQ:SE User's Guide". Baltimore: Paul H. Brookes Publishing.
- Tupes, E.C., y Christal, R.E., (1961). "Recurrent Personality Factors Based on Trait Ratings. Technical Report ASD-TR-61-97", Lackland Air Force Base, TX: Personnel Laboratory, Air Force Systems Command.
- Kochanska, G., Murray, K. T., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). "Inhibitory control in young children and its role in emerging internalization". *Child Development*, 67, 490–507



Departamento de Economía

Universidad de Chile