



Encuesta Longitudinal de Primera Infancia Cuarta Ronda 2024

Informe Metodológico

2025-10-30

División Observatorio Social
Subsecretaría de Evaluación Social
Ministerio de Desarrollo Social y Familia

Participaron en la elaboración de este documento: Cristina Marchant, Vicky Rojas, Gabriel Sotomayor y Diego Rodríguez, con el apoyo de las y los profesionales del equipo de la División Observatorio Social y la jefatura de división.

Tabla de contenidos

1	ANTECEDENTES CUARTA RONDA ELPI 2024	4
2	RECOLECCIÓN	6
2.1	CONFORMACIÓN DEL MARCO MUESTRAL Y SELECCIÓN DE LA MUESTRA	6
2.1.1	Población Objetivo	6
2.1.2	Marco Muestral	7
2.1.3	Evolución de la Muestra ELPI: 2010 a 2024	7
2.1.4	Diseño muestral ELPI 2010	7
2.1.5	Actualización de contactos	8
	2.1.5.1 Actualización de contactos 2019	8
	2.1.5.2 Actualización de contactos 2023	9
2.2	PREPARACIÓN DE LA RECOLECCIÓN	9
2.2.1	Capacitación	10
	2.2.1.1 Capacitación de coordinadores(as)	10
	2.2.1.2 Capacitación de encuestadores(as)	11
2.2.2	Difusión	12
2.3	EJECUCIÓN DE LA RECOLECCIÓN	14
	2.3.0.1 Estructura del equipo de terreno y flujo de aplicación	14
2.3.1	Logro de la muestra	15
2.3.2	Códigos de disposición final de casos	17
2.3.3	Tasas de respuesta	19
2.3.4	Atrición	20
2.3.5	Tiempos de aplicación	26
2.3.6	Supervisión del trabajo de campo	27

2.3.6.1	Supervisión Externa	28
2.3.6.2	Supervisión CMD	33
2.3.6.3	Supervisión MDSF	36
3	PROCESAMIENTO	38
3.1	REVISIÓN Y VALIDACIÓN DE DATOS	38
3.1.1	Validación de datos por el Centro de Microdatos	38
3.1.2	Validación de datos por el Ministerio de Desarrollo Social	40
3.1.2.1	Validaciones Posteriores (Post-Campo)	41
3.2	CLASIFICACIÓN Y CODIFICACIÓN	41
3.3	EDICIÓN DE DATOS	43
3.3.1	Edición de Variables	43
3.3.2	Imputación de Ingresos	44
3.4	CÁLCULO DE NUEVAS VARIABLES	46
3.5	CÁLCULO DE FACTORES DE EXPANSIÓN	47
3.6	ARCHIVOS FINALES DE DATOS	47

Capítulo 1

ANTECEDENTES CUARTA RONDA ELPI 2024

La Encuesta Longitudinal de Primera Infancia (ELPI) es un estudio que recoge datos en distintos momentos del tiempo sobre una muestra de niños y niñas¹, y sus responsables principales. La ELPI surge a partir de la inquietud de contar con datos respecto al desarrollo infantil en Chile, bajo un diseño de tipo longitudinal. De esta manera, la encuesta se desarrolló en su primera ronda del año 2010, al alero del Ministerio de Educación, que asumió este desafío, para pasar durante su segunda ronda del año 2012 al Ministerio del Trabajo y Previsión Social. En el año 2017, la ELPI fue asumida en su tercera ronda por el Ministerio de Desarrollo Social y Familia (MDSF), manteniéndose en esta institución hasta la fecha.

El objetivo de ELPI ha sido caracterizar y analizar el desarrollo de niños y niñas a lo largo de su nacimiento, infancia y adolescencia, considerando el efecto que pueden tener las características del hogar, de la principal persona responsable de su cuidado (en adelante, responsable principal) y del entorno cercano de niños y niñas sobre sus trayectorias de vida, con el propósito de informar la política pública, actual y futura, orientada a la infancia y adolescencia.

La población objetivo en la primera ronda de 2010 fueron los niños y niñas nacidas entre el 1 de enero de 2006 y el 31 de agosto de 2009 inscritos en el Servicio de Registro Civil e Identificación de Chile. El diseño muestral fue bietápico estratificado por conglomerados, seleccionando comunas en una primera etapa, y niños y niñas en la segunda etapa. Durante la primera ronda, se logró entrevistar un total de 15.175 casos considerando a los niños y niñas, y sus responsables principales.

¹El Ministerio de Desarrollo Social y Familia adhiere al enfoque de género utilizando un lenguaje inclusivo. De esta manera, en el documento nos referiremos a niños, niñas y adolescentes, evitando la utilización de siglas que puedan invisibilizarles, o reemplazando por niñez y adolescencia cuando sea necesario (<https://www.defensorianinez.cl/wp-content/uploads/2021/10/Minuta-Lenguaje-F%C3%A1cil.pdf>). Asimismo, se prioriza el uso de términos inclusivos, como “personas responsables”, en lugar de desdoblar sistemáticamente los sustantivos, adjetivos y artículos en sus versiones masculina y femenina. Este enfoque se aplicará en otras instancias similares a lo largo del texto. Sin embargo, en ciertos casos, para asegurar la claridad de la exposición, estas reglas generales podrían tener excepciones.

Para el año 2012 y 2017 se realizó un seguimiento a la muestra de 2010, agregando en cada versión una muestra de refresco² de nuevas cohortes de nacimiento; conformando lo que se conoce como panel con renovación de nacimientos³. Para la cuarta ronda se esperaba desarrollar un diseño similar a años anteriores, haciendo un seguimiento de la muestra de 2010, y las muestras de refresco de 2012 y 2017, incluyendo además una nueva cohorte de nacimiento. Sin embargo, durante los años 2021, 2022 y 2023 se levantaron cuatro procesos de licitación pública, que quedaron desiertos ya sea por falta de oferentes o porque las ofertas excedían el presupuesto.

Esto llevó a un replanteamiento del diseño de ELPI, con el fin de hacerlo viable. De esta manera, para la cuarta ronda se decidió realizar un seguimiento sólo de la muestra de 2010. Esto implicó elaborar una medición para una población exclusivamente de adolescentes y jóvenes entre 14 a 18 años y sus responsables principales, manteniendo a su vez una continuidad temática y metodológica con las rondas anteriores. Para llevar a cabo este proceso, el MDSF adjudicó, mediante licitación pública, el levantamiento de la encuesta al Centro de Microdatos (CMD) de la Facultad de Economía y Negocios de la Universidad de Chile. Adicionalmente, el diseño y ejecución del estudio contaron con la asesoría de un Panel de Expertos convocado para este fin y la colaboración del Instituto Nacional de Estadísticas (INE) en materias de construcción de factores de expansión.

En este documento, se reportan los procesos de recolección y procesamiento de la cuarta ronda de ELPI llevados a cabo entre mayo y octubre de 2024, bajo el marco del Generic Statistical Business Process Model (GSBPM), el cual describe una serie de procesos necesarios para la producción de estadísticas oficiales, proveyendo un marco estándar y una terminología armonizada (UNECE, 2019).

El Capítulo 2 de este documento describe el proceso de recolección de la encuesta, detallando la conformación del marco muestral, la preparación del trabajo de campo (incluyendo capacitación y difusión) y la ejecución de la recolección, donde se presentan los resultados de logro, tasas de respuesta, tiempos de aplicación, análisis de atrición y los distintos procesos de supervisión. Posteriormente, el Capítulo 3 aborda el procesamiento de los datos, abarcando la revisión y validación de la información, la codificación y clasificación de variables, la edición de datos y la imputación de ingresos, construcción de nuevas variables y de los factores de expansión.

²El diseño muestral se repitió para las muestras de refresco de las rondas siguientes. Se entiende como “muestra de refresco” la adición de nuevos casos dentro de una muestra previamente establecida en una encuesta de tipo longitudinal; esto puede obedecer a razones del impacto de la no respuesta o atrición en la representatividad de la muestra, a la necesidad de sobrerrepresentar un segmento de la población, o para cubrir nuevas “entradas” de la población (Watson & Lynn, 2021). En el caso de ELPI, obedece a la cobertura de nuevos nacimientos de niños y niñas, ya que se consideraron nuevas cohortes de nacimiento.

³Para diferencias en tipos de paneles ver Lynn (2005).

Capítulo 2

RECOLECCIÓN

Este capítulo describe en detalle el proceso de recolección de datos de la cuarta ronda de la Encuesta Longitudinal de Primera Infancia (ELPI 2024). La exposición se organiza en tres secciones principales que siguen el flujo cronológico del trabajo estadístico. En primer lugar, se aborda la conformación del marco muestral y la selección de la muestra, explicando la evolución histórica de las cohortes de la encuesta y detallando el diseño muestral que dio origen al panel. A continuación, se describe la fase de preparación de la recolección, que incluye los procesos de capacitación del equipo de terreno y las estrategias de difusión implementadas para facilitar el contacto con los participantes. Finalmente, se presenta la ejecución de la recolección, donde se reportan los resultados operativos del trabajo de campo, incluyendo el logro de encuestas, el análisis de los códigos de disposición final, las tasas de respuesta estandarizadas, los tiempos de aplicación, un análisis de la atrición muestral y un resumen de los distintos procesos de supervisión que se llevaron a cabo para garantizar la calidad de la información.

2.1 CONFORMACIÓN DEL MARCO MUESTRAL Y SELECCIÓN DE LA MUESTRA

2.1.1 Población Objetivo

La población objetivo de la cuarta ronda de ELPI corresponde a los niños, niñas y adolescentes nacidos entre el 1 de enero de 2006 y el 31 de agosto de 2009, que fueron inscritos en el Servicio de Registro Civil e Identificación (SRCel) y residen habitualmente en Chile.

Esta definición se enmarca en la decisión de dar continuidad a la muestra original de la encuesta. Por lo tanto, la población objetivo de 2024 es el resultado del seguimiento de la población objetivo definida en la primera ronda de ELPI, llevada a cabo en 2010.

2.1.2 Marco Muestral

Para la primera ronda de ELPI, el Marco Muestral (MM) se construyó a partir de los registros administrativos del SRCel. Correspondió a 877.084 registros de nacimientos ocurridos entre el 1º de enero de 2006 y el 31 de agosto de 2009. Este universo incluía a todos los niños y niñas nacidos vivos en dicho período y que no registraban defunción a la fecha de la construcción del marco.

Debido a que el marco se basa en el registro de nacimientos en Chile, se excluyó de la población objetivo a los niños y niñas nacidos en el extranjero de progenitores no chilenos. Adicionalmente, se aplicaron criterios de exclusión operativos, marginando a quienes residían en comunas de difícil acceso geográfico o en asentamientos que no contaban con una división administrativa formal de comuna, para asegurar la factibilidad del levantamiento en terreno.

2.1.3 Evolución de la Muestra ELPI: 2010 a 2024

El diseño de la Encuesta Longitudinal de Primera Infancia (ELPI) se ha construido a lo largo de varias rondas, comenzando con una cohorte fundacional en 2010 y expandiéndose en rondas posteriores para seguir el desarrollo de nuevas generaciones.

- Ronda 2010 (Muestra Original): La primera ronda estableció el panel inicial, seleccionando y logrando una muestra representativa de 15,175 niños y niñas de la población objetivo. Esta cohorte constituye el núcleo del seguimiento longitudinal de la encuesta.
- Ronda 2012 (Primer Refresco): En la segunda ronda, además de seguir a la cohorte original, se añadió una muestra de refresco de 3,135 niños y niñas (nacidos entre septiembre de 2009 y diciembre de 2011) para ampliar la cobertura a nuevas cohortes.
- Ronda 2017 (Segundo Refresco): La tercera ronda continuó el seguimiento de las muestras de 2010 y 2012, e incorporó una segunda muestra de refresco de 4,935 niños y niñas (nacidos entre enero de 2012 y diciembre de 2016).
- Ronda 2024 (Seguimiento Exclusivo): Debido a dificultades en los procesos de licitación, la cuarta ronda de ELPI se focalizó exclusivamente en realizar el seguimiento de la muestra original de 2010. Por lo tanto, el universo de la encuesta para esta ronda correspondió a los 15,175 participantes de la cohorte fundacional, quienes al momento del levantamiento eran adolescentes y jóvenes de entre 14 y 18 años.

2.1.4 Diseño muestral ELPI 2010

La ELPI se encuentra diseñada para obtener indicadores confiables con errores de muestreo aceptables para los niveles nacional, regional y según tramos de edad de

los niños y niñas dependiendo de la ronda.

El diseño muestral de la ELPI es un diseño probabilístico bietápico. La primera etapa corresponde a la selección de comunas donde, con probabilidad igual a 1 se incluyen 83 comunas que corresponden a todo el Gran Santiago, Coyhaique y, todas aquellas que tengan más de 60.000 habitantes, según el Censo 2002. Las restantes 263 comunas se agrupan en 33 estratos muestrales. Para ello, se ordenan según región, ingreso per cápita y población de niños y niñas según la información de Casen 2006 formando conjuntos de 100.000 habitantes aproximadamente en cada estrato. En cada uno de estos grupos se calcula la probabilidad proporcional al tamaño (PPT) para cada comuna, luego según esta probabilidad y el método de imagen inversa, se selecciona una comuna que represente al conjunto completo. Según lo anterior, la muestra contiene 116 comunas: 83 de inclusión automática y 33 seleccionadas desde los grupos de comunas creados.

En la segunda etapa de selección, considerando la selección en la etapa anterior y la información recolectada en la encuesta Casen del año correspondiente, además de los datos proporcionados por el SRCel, dentro de cada comuna se ordenan los registros de los niños y niñas de acuerdo con su sexo y edad, luego se selecciona mediante un muestreo sistemático, a partir del salto sistemático calculado como $k = \frac{N_{ch}}{n}$, donde N es la población de niños y niñas en la comuna seleccionada y n corresponde al número de niños y niñas a seleccionar en la muestra de la comuna seleccionada.

2.1.5 Actualización de contactos

Entre la tercera ronda de 2017 y la cuarta ronda de 2024, se realizaron dos procesos de verificación y actualización de los contactos de los niños, niñas y adolescentes (NNA) de la muestra, con el objetivo de reducir la pérdida de contacto y fortalecer la vinculación de los participantes con el estudio. El primer proceso fue ejecutado en 2019 por el Centro de Microdatos de la Universidad de Chile, y el segundo por EES Ingeniería en 2023, ambos mediante modalidad telefónica. En los dos casos, la adjudicación se realizó a través de licitación pública.

2.1.5.1 Actualización de contactos 2019

El proceso de 2019, denominado “Verificación, Actualización y Fidelización de Muestra de la Encuesta Longitudinal de Primera Infancia (ELPI)”, incluyó a la muestra original de 2010 y las muestras de refresco de 2012 y 2017, totalizando 23.336 casos. De estos, se logró entrevistar a 11.817 participantes: 11.750 de manera completa y 67 de forma parcial, alcanzando una tasa de éxito del 50,35%. Durante las entrevistas, se actualizó información sobre nombres, números telefónicos, correos electrónicos y direcciones del cuidador principal, padres, madres y otros contactos asociados a los NNA.

Como complemento a las llamadas telefónicas, el 28 de noviembre se enviaron correos electrónicos para actualizar datos de 4.217 personas no localizadas. Adicionalmente, el

23 de diciembre de 2019 se distribuyó un mensaje navideño dirigido a los participantes contactados que disponían de correo electrónico registrado.

Posteriormente, las restricciones impuestas por la pandemia de COVID-19 entre 2020 y 2021 limitaron la posibilidad de realizar encuestas presenciales, lo que, sumado a las dificultades para adjudicar nuevas licitaciones, provocó un retraso significativo en el trabajo de campo de la cuarta ronda de la ELPI.

2.1.5.2 Actualización de contactos 2023

En este contexto, a principios de 2023 se implementó un nuevo proceso de verificación y actualización centrado exclusivamente en la muestra original de 2010, compuesta por 15.175 casos. El objetivo principal fue contactar y actualizar los datos necesarios para facilitar la implementación de la cuarta ronda de la ELPI. Para ello, se utilizaron como base los datos recolectados en la ronda de 2017, complementados con los resultados de la actualización de 2019 y registros administrativos actualizados al año 2022, que incluían información sobre los cuidadores principales, madres, padres y los colegios de los niños, niñas y adolescentes.

El trabajo de terreno, que se desarrolló entre el 2 de enero y el 10 de febrero de 2023, involucró llamadas telefónicas realizadas desde un call center, el envío de correos electrónicos, mensajes de texto y el uso de redes sociales. Como resultado, se logró actualizar 5.299 folios, representando un 35% de la muestra inicial. Las llamadas telefónicas fueron el método más eficaz, logrando actualizar 5.119 casos. Por su parte, los correos electrónicos permitieron localizar a 180 participantes adicionales que no habían sido contactados por otras vías. Las redes sociales, aunque utilizadas como un recurso complementario, demostraron ser menos efectivas debido a restricciones en las plataformas y dificultades para establecer contacto directo.

El perfil de los contactos exitosos reflejó que las madres de los niños constituyeron el 68,3% de las personas localizadas, mientras que los padres representaron el 22,3%. Además, el 85% de los participantes contactados informó residir con el menor, y el 79,9% declaró ser el cuidador principal.

2.2 PREPARACIÓN DE LA RECOLECCIÓN

Como se mencionó, el levantamiento de la cuarta ronda de ELPI fue adjudicado mediante licitación pública al Centro de Microdatos (CMD) de la Universidad de Chile. Previo al inicio del trabajo de campo, se llevó a cabo un exhaustivo proceso de preparación que incluyó la capacitación del equipo de terreno y la difusión de la encuesta a la población objetivo y a actores clave. Estas tareas fueron diseñadas y ejecutadas de manera conjunta entre el CMD y la División de Observatorio Social (DOS) del MDSF.

2.2.1 Capacitación

El proceso de capacitación buscó garantizar la calidad y homogeneidad en la aplicación de los instrumentos y en la gestión del terreno. Se estructuró bajo un modelo de cascada en dos grandes etapas, que se describen en detalle a continuación. En la primera etapa, el equipo del MDSF y del Centro de Microdatos capacitó de forma centralizada en Santiago a las personas coordinadoras, quienes serían responsables de la gestión, logística y supervisión del levantamiento en cada una de las sedes regionales. En la segunda etapa, estos mismos coordinadores y coordinadoras replicaron la capacitación en sus respectivas regiones a las personas encuestadoras, encargadas de realizar las entrevistas en los hogares.

2.2.1.1 Capacitación de coordinadores(as)

La capacitación de las personas coordinadoras se desarrolló en tres etapas. La primera, realizada el 22 de abril de 2024, consistió en una jornada de autoformación de ocho horas. Los y las coordinadoras revisaron de manera autónoma el manual de trabajo de campo y los instrumentos de la encuesta a través de una plataforma interactiva diseñada por el CMD, que incluyó evaluaciones de conocimientos sobre los temas revisados. Este material de estudio se complementó con recursos digitales, lo que facilitó una comprensión previa de los contenidos antes de las actividades presenciales.

La segunda etapa se llevó a cabo entre el 24 y el 26 de abril de 2024 de manera presencial y centralizada en la Región Metropolitana. Durante tres días, los coordinadores y coordinadoras participaron en jornadas de capacitación de ocho horas cada una. Estas sesiones constaron de presentaciones de profesionales del MDSF y del CMD. Los contenidos abarcaron desde los antecedentes y objetivos del proyecto hasta la revisión detallada de los cuestionarios y evaluaciones, así como los protocolos de aplicación.

En el caso del Test de Vocabulario en Imágenes Peabody (TVIP), evaluación que busca medir vocabulario receptivo, y cuya aplicación requiere del uso de un cuadernillo de láminas con imágenes y conocimientos específicos sobre sus reglas de aplicación, se contó con una psicóloga experta que capacitó al equipo de coordinación en la administración del test. El CMD lideró ejercicios prácticos con simulaciones, en las cuales los participantes utilizaron tablets para poner en práctica los procedimientos.

La tercera etapa, desarrollada los días 15 y 16 de mayo de 2024, se realizó de manera online y abordó aspectos administrativos y de gestión de la muestra. Los y las coordinadoras revisaron la implementación de sedes, la planificación del calendario de actividades, la gestión de la muestra a través de plataformas digitales y otros temas logísticos clave para el desarrollo del proyecto. Esta fase complementó las actividades anteriores y permitió consolidar la preparación del equipo.

En total, participaron 30 coordinadores y coordinadoras provenientes de todas las regiones del país, cuya distribución puede observarse en la Tabla [2.1](#).

Tabla 2.1: Número de coordinadores(as) capacitados(as) por región

Región	Total asistentes
Arica y Parinacota	1
Tarapacá	1
Antofagasta	1
Atacama	1
Coquimbo	1
Valparaíso	3
Metropolitana	9
O'Higgins	2
Maule	2
Biobío	2
La Araucanía	2
Los Ríos	1
Los Lagos	2
Aysén	1
Magallanes	1
Total	30

2.2.1.2 Capacitación de encuestadores(as)

Las capacitaciones a encuestadores y encuestadoras fueron realizadas por los equipos de coordinación en cada región, entre las semanas del 13 y 27 de mayo de 2024, con un total de 21 capacitaciones y 285 personas encuestadoras capacitadas. En la tabla Tabla 2.2 puede observarse la distribución por región de las capacitaciones realizadas y el total de personas capacitadas.

Tabla 2.2: Número de encuestadores(as) capacitados(as) por región

Región	Total de capacitaciones	Total Asistentes
Arica y Parinacota	1	6
Tarapacá	1	9
Antofagasta	1	17
Atacama	1	9
Coquimbo	1	15
Valparaíso	1	23
Metropolitana	4	93
O'Higgins	1	14
Maule	1	13
Ñuble	1	9
Biobío	2	30
Araucanía	1	16
Los Ríos	1	5
Los Lagos	2	19
Aysén	1	3
Magallanes	1	4
Total	21	285

Las capacitaciones iniciaron con una jornada de autoformación no presencial, durante la cual los participantes revisaron, a través de una plataforma online, el manual del encuestador y realizaron pruebas de conocimiento para evaluar la comprensión de los contenidos.

Posteriormente, se llevaron a cabo jornadas presenciales durante los tres días siguientes, con una duración aproximada de ocho horas cada una. En estas sesiones se profundizó en el análisis de los cuestionarios y evaluaciones de ELPI, los protocolos de entrevista, el uso de consentimientos y asentimientos, así como en las metodologías de aplicación de instrumentos. Los y las asistentes participaron en ejercicios prácticos con tablets, que incluyeron simulaciones de entrevistas para resolver dudas en tiempo real. Se prestó especial atención a la capacitación en el uso del test TVIP, apoyada por cápsulas audiovisuales y un guion estructurado para los ejercicios. El tercer día de capacitación presencial concluyó con una prueba de contenidos para evaluar el nivel de aprendizaje de las y los encuestadores.

Con el objetivo de estandarizar los contenidos en cada región, los equipos de coordinación dispusieron de presentaciones uniformes para cada módulo, además de materiales audiovisuales centrados en los temas de mayor relevancia y/o complejidad. Entre estos se incluyó la descripción general de la encuesta y sus objetivos, el protocolo de incidencias y situaciones críticas, y la aplicación del TVIP. Cabe destacar que este último fue impartido por la misma psicóloga experta que colaboró en la capacitación de coordinadores. Adicionalmente, se realizó una capacitación online con los encuestadores y encuestadoras para reforzar esta materia.

2.2.2 Difusión

Para facilitar el levantamiento de la cuarta ronda de la Encuesta Longitudinal de Primera Infancia (ELPI), se implementaron diversas estrategias comunicacionales diseñadas y ejecutadas en conjunto por el Ministerio de Desarrollo Social y Familia y el Centro de Microdatos. Estas acciones tuvieron como objetivo principal dar a conocer ampliamente el propósito y alcance de la encuesta, además de asegurar un adecuado respaldo institucional durante la realización del trabajo de campo.

Entre las actividades más relevantes destacó la organización de un lanzamiento oficial mediante el seminario titulado «Distintas perspectivas sobre el aporte de la Encuesta ELPI», realizado en dependencias de la Facultad de Economía y Negocios de la Universidad de Chile. Este evento contó con la participación de autoridades del MDSF, representantes de UNICEF Chile y de la Defensoría de la Niñez, así como académicos de la misma facultad. La instancia permitió resaltar públicamente la importancia de la ELPI como herramienta fundamental para monitorear el desarrollo infantil en Chile y generar evidencia clave para la formulación de políticas públicas orientadas a reducir desigualdades desde la primera infancia. A partir de este seminario, se generaron diversos productos comunicacionales para ampliar su difusión, tales como comunicados de prensa, un video informativo y piezas gráficas para redes sociales.

Adicionalmente, se elaboró material específico dirigido directamente a las personas entrevistadas. Este consistió en un tríptico/carta informativo que detallaba los objetivos del estudio, las características de la participación y los beneficios de formar parte de la encuesta.



Figura 2.1: Imagen del tríptico-carta informativo para participantes.



Figura 2.2: Imagen del tríptico-carta informativo para participantes.

Para reforzar la transparencia y el acceso a la información, se crearon dos plataformas web. Por un lado, el sitio web institucional del CMD difundió detalles técnicos y operativos

del levantamiento. Por otro, el Ministerio de Desarrollo Social y Familia habilitó una página web (observatorio.ministeriodesarrollosocial.gob.cl/elpi-cuarta-ronda) con información general sobre la encuesta, su historia y los objetivos de la ronda actual. Este sitio sirvió como respaldo y fuente de consulta tanto para los participantes como para el equipo de encuestadores(as) en terreno.

Para apoyar la labor de los equipos de campo, se emitió desde el MDSF una carta informativa dirigida a Carabineros de Chile. Este documento impreso fue portado por los encuestadores y encuestadoras como un respaldo institucional formal frente a situaciones imprevistas que pudiesen presentarse durante la realización de las entrevistas.

Asimismo, se gestionó un oficio dirigido al Ministerio de Educación (MINEDUC), solicitando su colaboración activa en la difusión de la encuesta en aquellos establecimientos educacionales donde se encontraban matriculados los y las adolescentes de la muestra.

Finalmente, en comunas con una alta concentración de la muestra o donde se identificaron mayores tasas de rechazo durante el levantamiento, se realizaron contactos específicos con autoridades municipales y con Carabineros, con el propósito de asegurar el adecuado desarrollo del levantamiento en estos territorios.

Estas iniciativas permitieron reforzar la legitimidad y visibilidad pública de la ELPI y el proceso de recolección, promoviendo condiciones más favorables para esta última.

2.3 EJECUCIÓN DE LA RECOLECCIÓN

La fase de recolección de datos de la cuarta ronda de la Encuesta Longitudinal de Primera Infancia (ELPI 2024) fue llevada a cabo por el Centro de Microdatos (CMD) de la Facultad de Economía y Negocios de la Universidad de Chile. El trabajo de campo se desarrolló en modalidad presencial, asistida por dispositivos móviles (CAPI), entre el **25 de mayo y el 15 de octubre de 2024**.

2.3.0.1 Estructura del equipo de terreno y flujo de aplicación

El equipo de terreno se organizó a nivel nacional con una estructura jerárquica para asegurar una gestión y supervisión efectivas. El equipo estuvo compuesto por 30 coordinadores(as) regionales, quienes lideraban el trabajo en cada sede, y 285 encuestadores(as), responsables directos de la aplicación de los instrumentos en los hogares.

La aplicación de la encuesta se diseñó para ser administrada mediante dos tablets por cada encuestador(a), con el fin de optimizar los tiempos y permitir la aplicación simultánea de instrumentos. El software utilizado para la programación y captura de datos fue *Survey Solutions*. El flujo de aplicación en el hogar fue el siguiente:

Paso	Dispositivo	Instrumento Aplicado	Respondente	Modalidad
1	Tablet 1	Cuestionario al Responsable Principal	Responsable Principal	Entrevista CAPI
2	Tablet 1	Cuestionario al Adolescente	Adolescente	Autoaplicado (CAPI)
3	Tablet 2	Evaluaciones al Responsable Principal	Responsable Principal	Autoaplicado (CAPI)
4	Tablet 1	Test de Vocabulario en Imágenes (TVIP)	Adolescente	Entrevista CAPI

Como muestra el esquema, la entrevista comenzaba con el cuestionario aplicado por el/la encuestador(a) al responsable principal en la primera tablet. Luego, en esa misma tablet, se le entregaba al adolescente para que respondiera su cuestionario de forma autoaplicada. De manera paralela a esto, el/la encuestador(a) entregaba la segunda tablet al responsable principal para que completara sus evaluaciones (escalas), también de forma autoaplicada. Finalmente, el/la encuestador(a) retomaba la primera tablet para aplicar directamente la evaluación TVIP al adolescente.

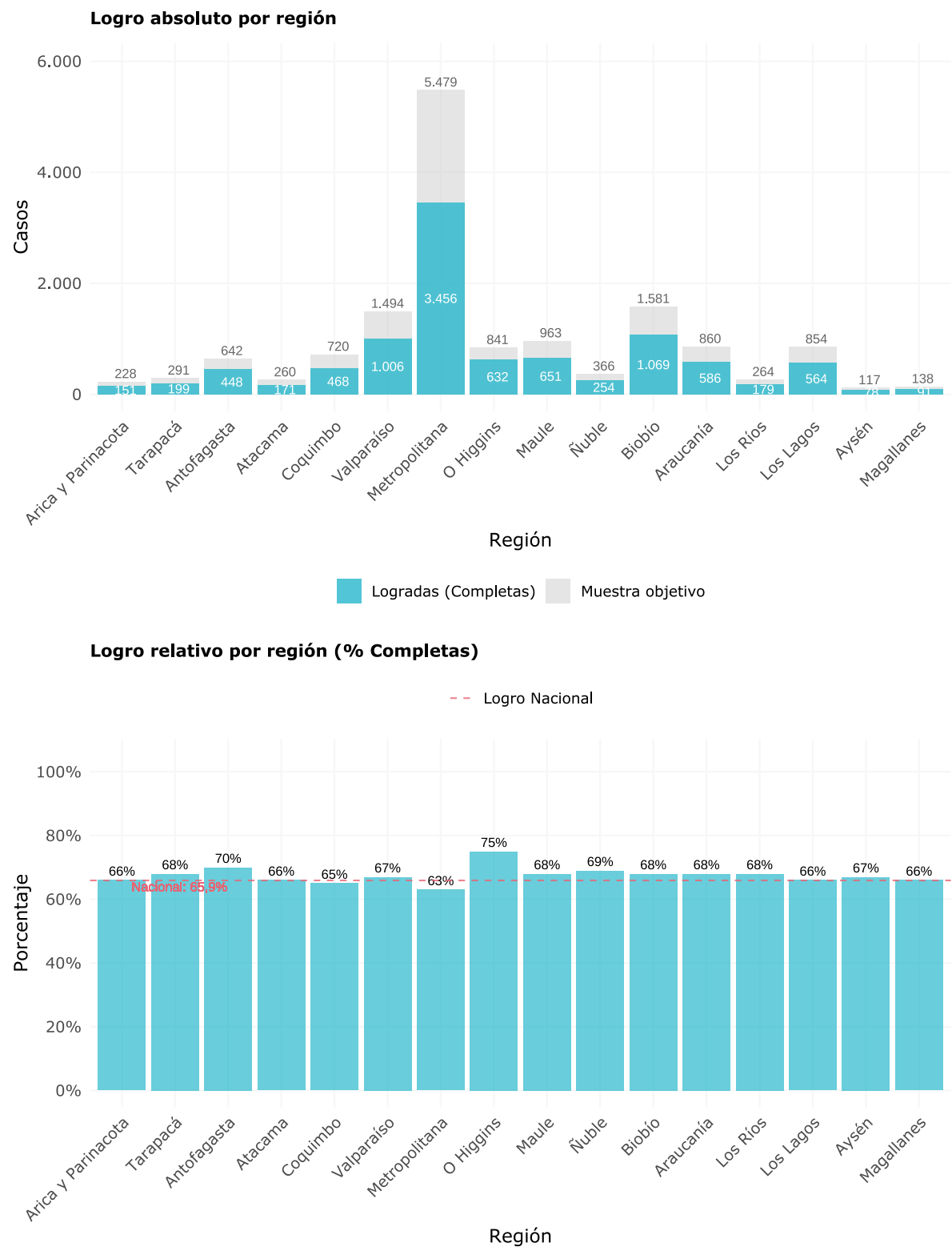
2.3.1 Logro de la muestra

De la muestra objetivo de 15.175 casos, correspondiente al total de la muestra lograda en ELPI 2010, se obtuvieron **10.003 encuestas completas**¹, lo que representa una **tasa de respuesta nacional del 65,9%** respecto a la muestra inicial. A nivel regional, se observa una distribución relativamente homogénea del logro, con la mayoría de las regiones alcanzando tasas de logro en torno al 65-70%. La Región de O'Higgins destaca con la tasa de logro más alta (75%), mientras que la Región Metropolitana presenta el porcentaje más bajo (63%), lo cual puede atribuirse a las mayores dificultades de contacto y las dificultades de gestión de una muestra considerablemente mayor a la de otras regiones.

La Figura 2.3 muestra el detalle del logro absoluto (número de encuestas) y relativo (porcentaje) para cada región del país.

¹Se considera "encuesta completa" aquella en la que se aplicaron exitosamente todos los instrumentos y evaluaciones tanto al responsable principal como al adolescente, según el flujo de aplicación definido.

Figura 2.3: Logro de terreno absoluto y relativo por región, ELPI 2024



Fuente: Elaboración propia en base a datos ELPI 2024 (DFC Final).

2.3.2 Códigos de disposición final de casos

Para clasificar el resultado de cada intento de contacto y entrevista, se utilizó un sistema estandarizado de Códigos de Disposición Final (CDF), basado en las recomendaciones de la *American Association for Public Opinion Research* (AAPOR). Estos códigos permiten distinguir entre encuestas completas, parciales, rechazos, casos de no contacto, otros motivos de no entrevista en casos elegibles, y casos donde la elegibilidad no pudo ser determinada.

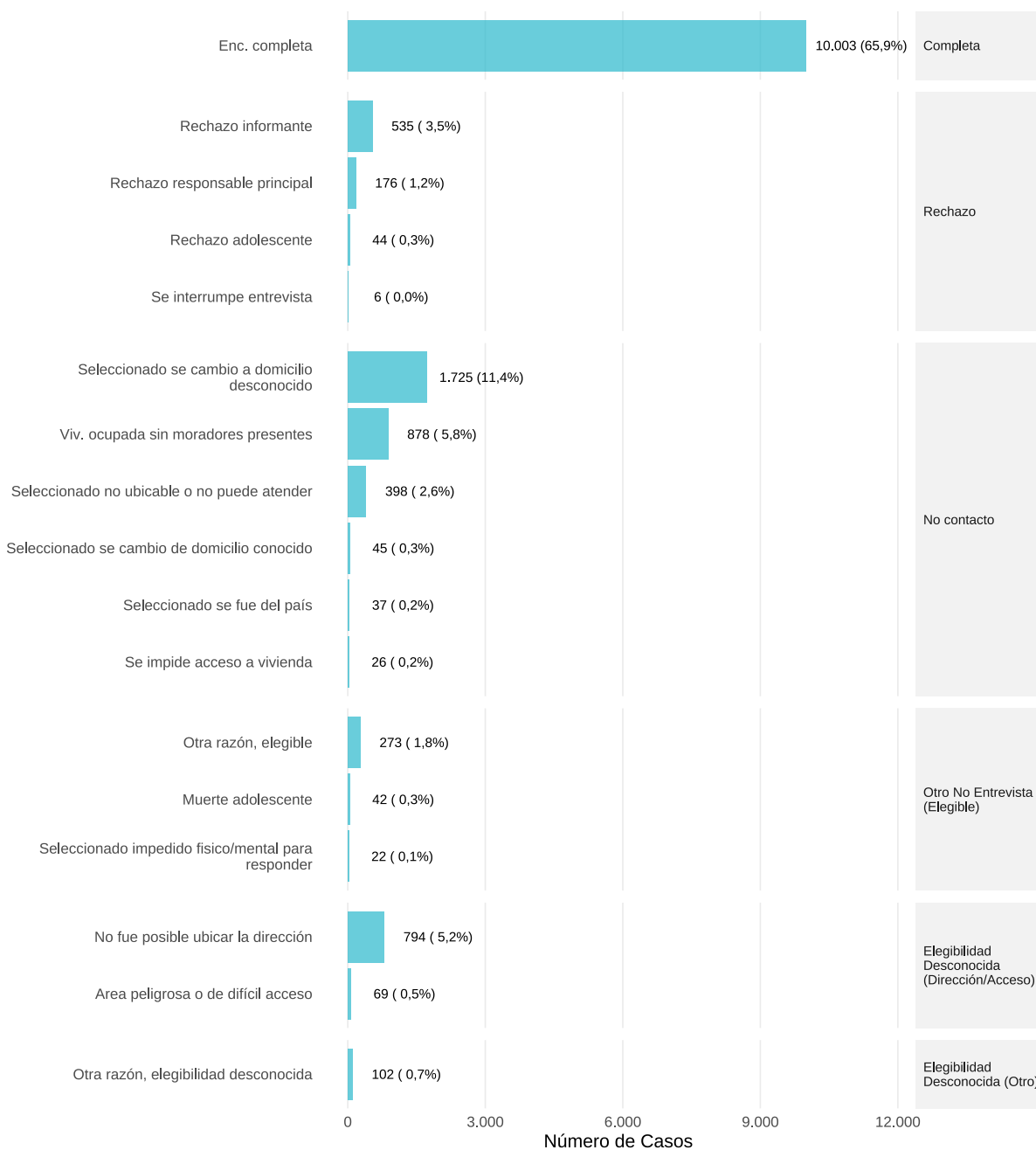
La siguiente tabla presenta el listado completo de los códigos de disposición final utilizados en la cuarta ronda de ELPI.

Tabla 2.4: Códigos de Disposición Final (CDF) utilizados en ELPI 2024

Código	Descripción
Completa	
110	Encuesta completa
Parcial	
120	Encuesta parcial (genérica)
121	Encuesta parcial a responsable principal
122	Encuesta parcial a adolescente
Rechazo	
211	Rechazo de entrevista por informante del hogar
2112	Rechazo de entrevista por parte del adolescente
21111	Rechazo de entrevista por parte del responsable principal
212	Se interrumpió la entrevista
No contacto	
223	Se impidió el acceso a la vivienda
224	Vivienda ocupada sin moradores presentes
225-0	Seleccionado/a no ubicable o no puede atender
225-1	Seleccionado/a se cambió a domicilio conocido
225-2	Seleccionado/a se cambió a domicilio desconocido
225-3	Seleccionado/a se fue del país
Otro No Entrevista (Elegible)	
231	Muerte del adolescente seleccionado
232	Seleccionado/a impedido/a físico/mental para responder
290	Otra razón, elegible
Elegibilidad Desconocida (Dirección Acceso)	
317	Área peligrosa o de difícil acceso
318	No fue posible localizar la dirección
Elegibilidad Desconocida (otro)	
390	Otra razón, elegibilidad desconocida
No entrevistados (temporal/administrativo)	
200	Agenda cita
510	En terreno (sin resultado final)
520	No asignada a terreno

La distribución final de los casos según las categorías agregadas de estos códigos se presenta en la Figura 2.4. Además de las 10.003 encuestas completas (C: 65,9%), se registraron 0 encuestas parciales (P: 0%). Los principales motivos de no respuesta incluyeron 761 rechazos (R: 5%), 3.109 casos de no contacto (NC: 20,5%), y 337 otros casos elegibles no entrevistados (O: 2,2%). Un total de 965 casos quedaron con elegibilidad desconocida (ED/OED: 6,4%), principalmente por dificultades para localizar la dirección o por encontrarse en áreas de difícil acceso. Dentro de los rechazos, el motivo más frecuente fue el rechazo explícito por parte del informante o del responsable principal. En los casos de no contacto, el cambio a un domicilio desconocido fue una de las causas más relevantes.

Figura 2.4: Distribución detallada de Códigos de Disposición Final por categoría, ELPI 2024



Fuente: Elaboración propia en base a datos ELPI 2024 (DFC Final).

2.3.3 Tasas de respuesta

Para evaluar el desempeño del levantamiento en términos de respuesta, se calcularon tasas estandarizadas siguiendo las definiciones de AAPOR. Las tasas principales consideradas son:

- **Tasa de Respuesta (RR2):** Proporción de encuestas completas y parciales sobre todos los casos elegibles más todos los casos de elegibilidad desconocida, asumiendo que los desconocidos son elegibles.

$$RR2 = \frac{C + P}{(C + P) + (R + NC + O) + (ED + OED)}$$

- **Tasa de Cooperación (COOP1):** Proporción de entrevistas completas sobre los casos contactados y elegibles donde se solicitó la entrevista. Mide la aquiescencia una vez establecido el contacto.

$$COOP1 = \frac{C}{(C + P) + R + O}$$

- **Tasa de Rechazo (REF1):** Proporción de rechazos sobre todos los casos elegibles o potencialmente elegibles.

$$REF1 = \frac{R}{(C + P) + (R + NC + O) + (ED + OED)}$$

- **Tasa de Contacto (CON1):** Proporción de todos los casos elegibles donde se logró establecer contacto (resultando en entrevista completa, parcial, rechazo u otro no entrevista elegible) sobre el total de casos elegibles o potencialmente elegibles.

$$CON1 = \frac{C + P + R + O}{(C + P) + (R + NC + O) + (ED + OED)}$$

A nivel nacional, la Tasa de Respuesta (RR2) final fue de 65,9%. La Tasa de Cooperación (COOP1) alcanzó un 90,1%, indicando una buena disposición a participar una vez que se lograba el contacto efectivo con un caso elegible. La Tasa de Rechazo (REF1) se situó en 5%, mientras que la Tasa de Contacto (CON1) fue de 73,2%. La Tabla 2.5 presenta el desglose de estas tasas por región.

Tabla 2.5: Tasas de respuesta, cooperación, rechazo y contacto por región (%), ELPI 2024

Región	Tasa Respuesta (RR2)	Tasa Cooperación (COOP1)	Tasa Rechazo (REF1)	Tasa Contacto (CON1)
Arica y Parinacota	66,2	90,4	7,0	73,2

Tarapacá	68,4	93,0	4,8	73,5
Antofagasta	69,8	98,2	1,1	71,0
Atacama	65,8	94,0	4,2	70,0
Coquimbo	65,0	95,1	3,1	68,3
Valparaíso	67,3	92,5	5,4	72,8
Metropolitana	63,1	84,9	6,4	74,3
O Higgins	75,1	97,2	2,1	77,3
Maule	67,6	91,3	4,5	74,0
Ñuble	69,4	93,4	4,4	74,3
Biobío	67,6	93,6	4,2	72,2
Araucanía	68,1	92,4	5,3	73,7
Los Ríos	67,8	96,2	2,7	70,5
Los Lagos	66,0	91,7	5,7	72,0
Aysén	66,7	95,1	3,4	70,1
Magallanes	65,9	87,5	9,4	75,4
Total Nacional	65,9	90,1	5,0	73,2

2.3.4 Atrición

Siendo ELPI una encuesta longitudinal que sigue a la misma cohorte de individuos a lo largo del tiempo, la **atrición** –la pérdida de participantes entre rondas sucesivas– es un fenómeno inherente y de suma importancia a analizar. La cuarta ronda de 2024 se centró en el seguimiento de la cohorte original de 15.175 niños y niñas (ahora adolescentes) reclutados en 2010.

Comparando la muestra inicial de 2010 con las 10.003 encuestas completas logradas en 2024, la tasa de **retención global** después de 14 años es del 65,9%. Complementariamente, la tasa de **atrición acumulada** (pérdida total desde 2010) asciende al 34,1%.

Las causas directas de la no participación en 2024 (reflejadas en los CDF distintos a “Completa”) incluyen el no contacto, rechazos, y otras razones como migración o fallecimiento. Un análisis más profundo permite examinar cómo la participación en rondas anteriores, así como características basales como el sexo o el año de nacimiento, influyen en la probabilidad de ser contactado y participar en la ronda actual (fenómeno conocido como **atrición diferencial**).

La Tabla 2.6 presenta la tasa de logro en 2024 según el historial de participación en las rondas 2010, 2012 y 2017. Como se observa, existe una fuerte relación entre la participación previa y la retención en 2024: aquellos que participaron en las tres rondas anteriores tuvieron una tasa de logro del 76,3%, mientras que quienes solo participaron en 2010 alcanzaron apenas un 35,2%. Esto subraya la importancia de la participación reciente para la continuidad en el estudio.

Tabla 2.6: Tasa de logro en 2024 según patrón de participación en rondas anteriores (2010, 2012, 2017)

Patrón Participación (Previo a 2024)	N Casos en Grupo	N Logrados 2024	Tasa Logro 2024 (%)
1. Solo 2010	1.243	437	35,2
2. 2010 + 2012	3.702	1.898	51,3

3. 2010 + 2017	1.034	656	63,4
4. 2010 + 2012 + 2017	9.196	7.012	76,3
Total	15.175	10.003	65,9

La Figura 2.5 visualiza las secuencias completas de participación a lo largo de las cuatro rondas, agrupando los casos por el número total de rondas participadas y ordenándolos por su patrón específico para mostrar las distintas trayectorias.

La Figura 2.6 desagrega el resultado final obtenido en 2024 según el patrón de participación *histórica* (hasta 2017), mostrando que la dificultad en el contacto es la principal causa de no respuesta para los grupos con menor participación previa.

Adicionalmente, se examinó la relación entre la tasa de logro en 2024 y el sexo del participante registrado en 2010, así como su año de nacimiento. Las tablas Tabla 2.7 y Tabla 2.8 resumen estas tasas.

Tabla 2.7: Tasa de logro en 2024 según sexo registrado en 2010

Sexo (2010)	N Casos en Grupo	N Logrados 2024	Tasa Logro 2024 (%)
Hombre	7.683	5.041	65,6
Mujer	7.492	4.962	66,2
Total	15.175	10.003	65,9

Tabla 2.8: Tasa de logro en 2024 según año de nacimiento

Año Nacimiento	N Casos en Grupo	N Logrados 2024	Tasa Logro 2024 (%)
2006	3.998	2.617	65,5
2007	4.096	2.715	66,3
2008	4.174	2.760	66,1
2009	2.906	1.911	65,8
Total	15.174	10.003	65,9

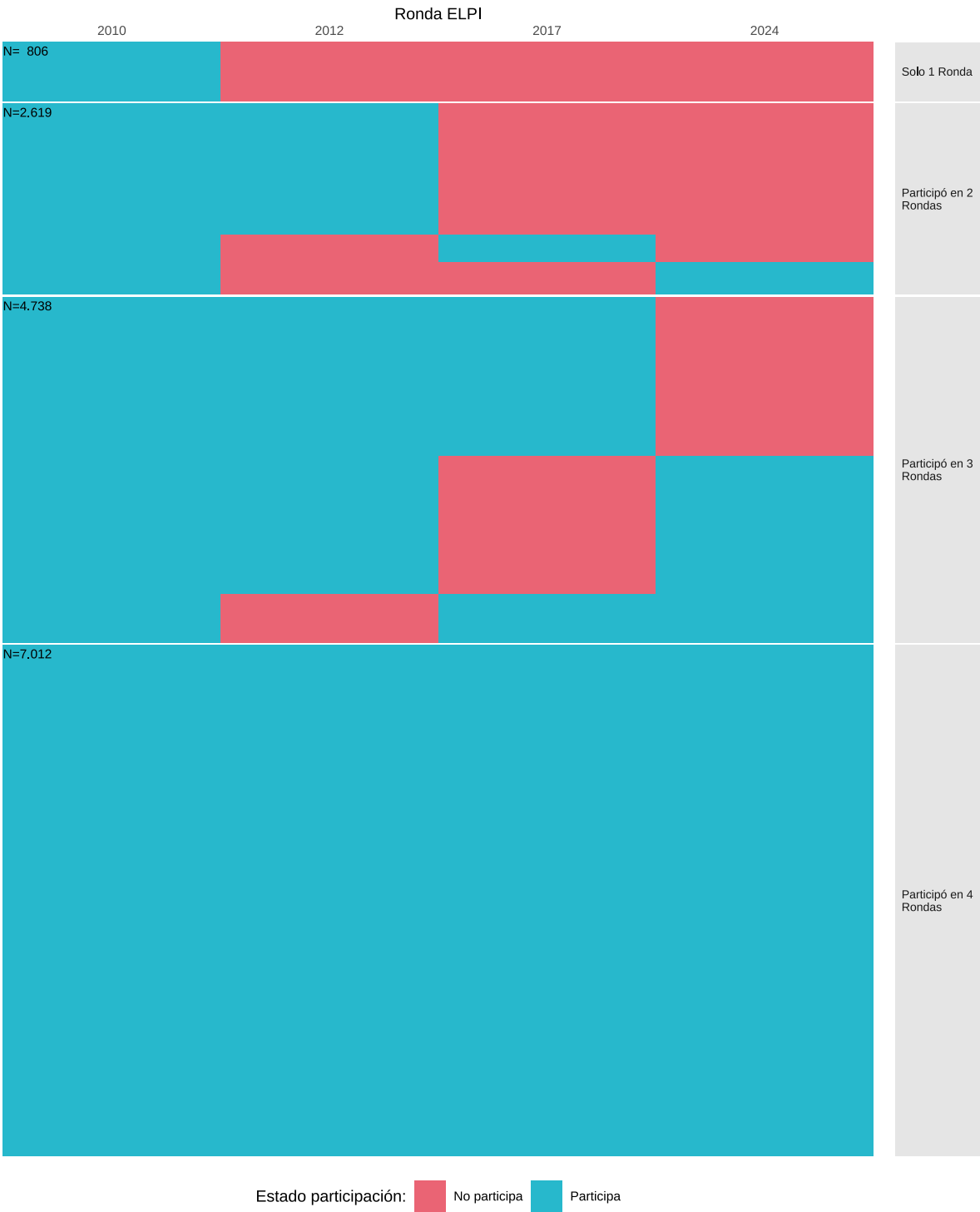
Como se aprecia en las tablas, las diferencias en la tasa de logro final según sexo (66,2% mujeres vs 65,6% hombres) y año de nacimiento (fluctuando entre 65,5% y 66,3%) son mínimas.

Las figuras Figura 2.7 y Figura 2.8, que muestran la distribución de los resultados finales según estas variables, refuerzan esta idea.

La composición de los resultados finales (Logrado, Rechazo, No contacto, Elegibilidad desconocida) es muy similar entre hombres y mujeres, al igual que según año de nacimiento.

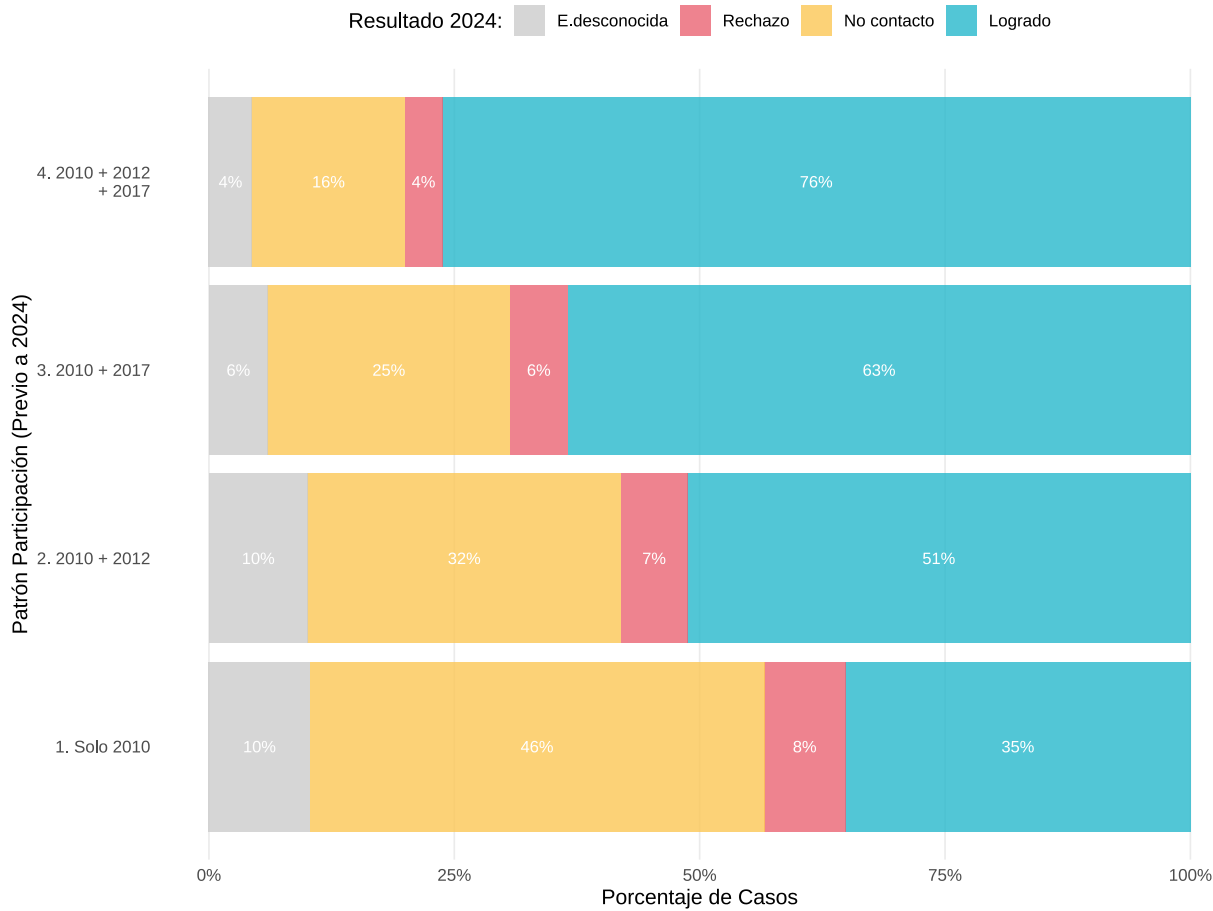
En conclusión, si bien la historia de participación previa emerge como el factor más fuertemente asociado a la probabilidad de completar la encuesta en 2024, la atrición parece ser relativamente homogénea con respecto al sexo y, en gran medida, al año de nacimiento de los participantes. La dificultad en el contacto es una causa importante de no respuesta, especialmente para aquellos casos con participaciones más lejanas en

Figura 2.5: Visualización de secuencias de participación en todas las rondas ELPI (2010-2024), agrupadas por número total de rondas participadas y ordenadas por patrón



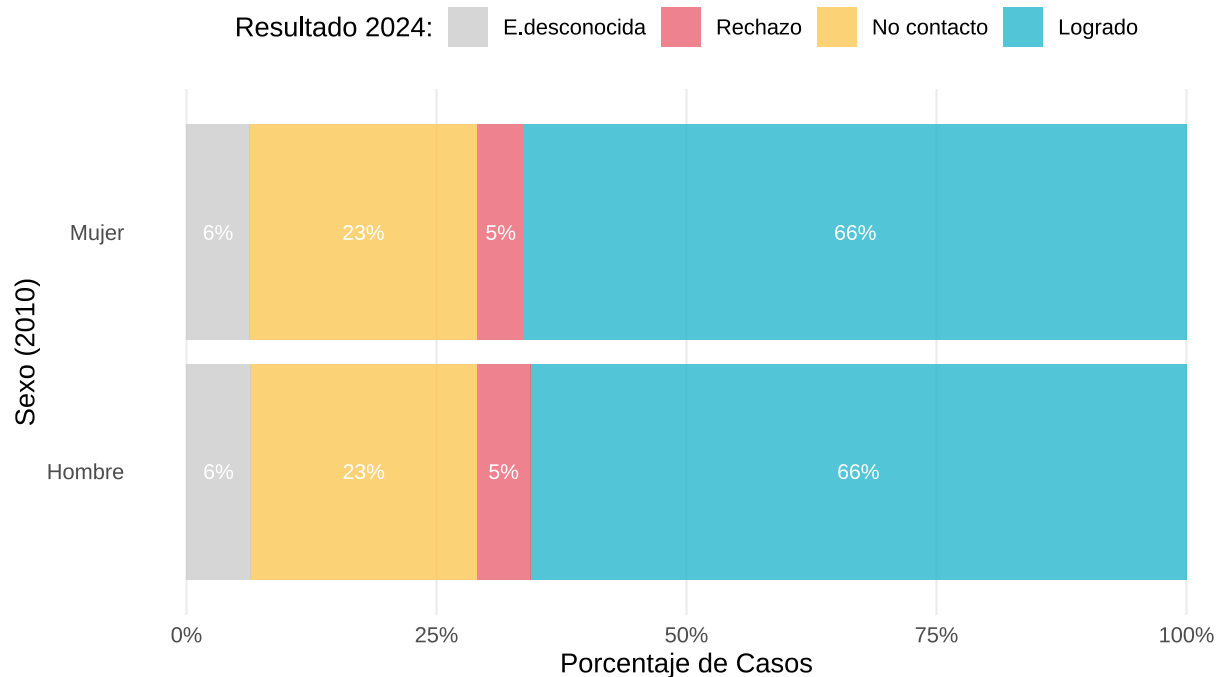
Fuente: Elaboración propia en base a datos ELPI 2024.
Cada línea representa un caso. Casos agrupados por N° total de rondas y ordenados por secuencia.

Figura 2.6: Distribución del resultado final en 2024 según patrón de participación histórica (hasta 2017)



Fuente: Elaboración propia en base a datos ELPI 2024.

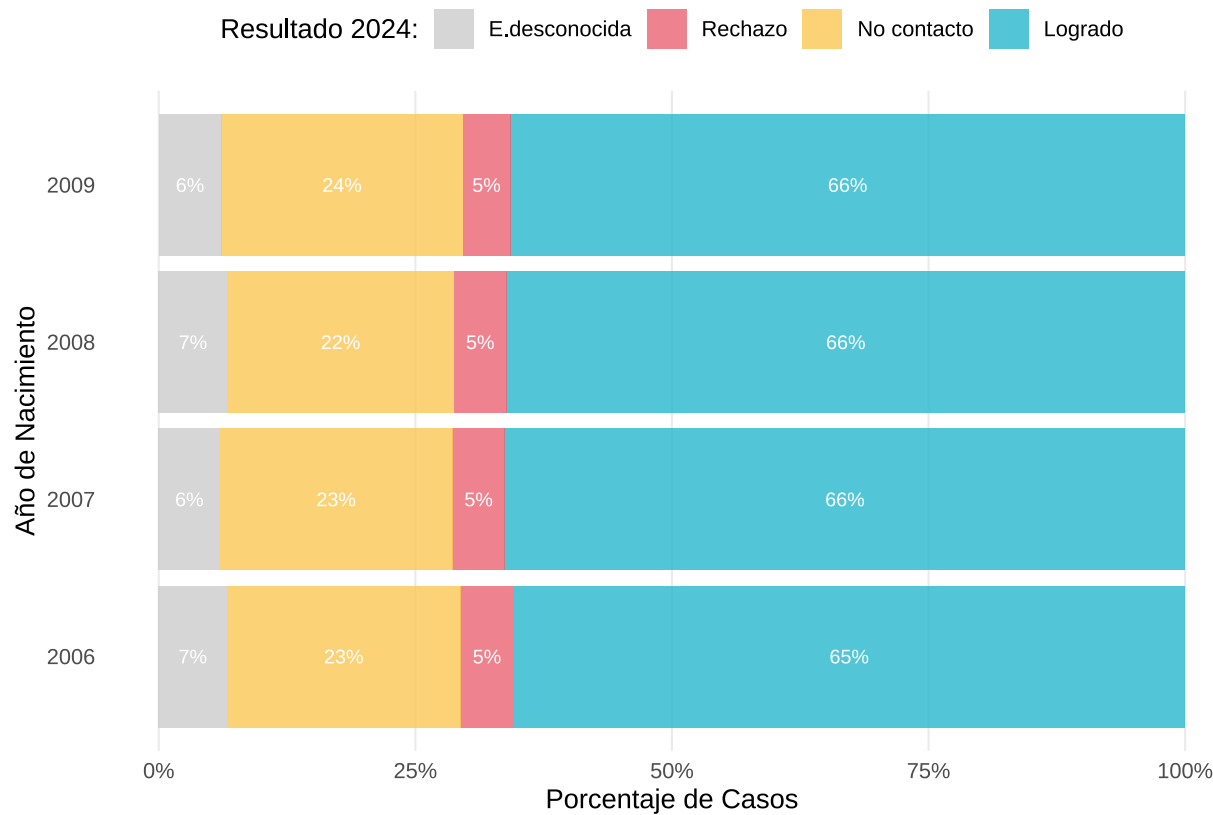
Figura 2.7: Distribución del resultado final en 2024 según sexo (2010)



el tiempo. Estos patrones fueron considerados en la etapa de ponderación para mitigar posibles sesgos ².

²Para mayores detalles sobre la construcción de los factores de expansión ver la sección 3.5 Cálculo de Factores de Expansión del presente informe.

Figura 2.8: Distribución del resultado final en 2024 según año de nacimiento



Fuente: Elaboración propia en base a datos ELPI 2024.

2.3.5 Tiempos de aplicación

La duración de la aplicación de la encuesta ELPI 2024 se analizó utilizando los registros de tiempo capturados por el sistema CAPI para sus distintos componentes: el cuestionario al responsable principal y adolescente, los tests autoaplicados al responsable y el Test de Vocabulario en Imágenes Peabody (TVIP). Como se detalló anteriormente, el uso de dos tablets permitió que el bloque de tests autoaplicados al responsable principal se realizara en paralelo mientras el o la adolescente respondía su propio cuestionario.

Para fines de análisis y comparabilidad, se calculó un tiempo total de aplicación lineal, que corresponde a la suma de los tres componentes. Sin embargo, también se presenta una estimación del tiempo real promedio de la entrevista (denominado “Tiempo Total Estimado (en paralelo)”), descontando el componente que se aplicaba de forma simultánea.

El análisis se basa en las 10.003 encuestas consideradas completas que contaban con registros de tiempo válidos para todos los componentes. Para mitigar la influencia de posibles valores extremos en la duración total, se aplicó un *trimming* bilateral a la distribución del tiempo total calculado, eliminando el 1,5% inferior y el 1,5% superior de los casos. Esto resultó en la exclusión de 302 casos adicionales, quedando 9.701 casos para el análisis final de tiempos.

La Tabla 2.9 presenta un resumen de las estadísticas descriptivas para el tiempo total lineal, sus componentes y el tiempo total estimado en paralelo a nivel nacional, considerando los datos trimados.

Tabla 2.9: Estadísticas descriptivas de tiempos de aplicación (minutos), Total Nacional (N=9.701 casos trimados al 3%)

	Promedio	Mediana	Desv. Estándar	Mínimo	Máximo
Tiempo Total Estimado (en paralelo)	113,3	106,4	34,8	38,6	262,5
Total Aplicación (suma lineal)	147,6	140,4	39,8	81,0	289,6
Cuestionarios RP+Adolescente	100,0	93,6	32,0	19,7	248,3
Tests Responsable Principal (TRP)	34,2	31,4	13,4	0,6	148,6
Test Vocabulario (TVIP)	13,3	10,8	8,4	0,4	117,7

Como se observa en la tabla anterior, el tiempo total estimado de aplicación tuvo un promedio de 113.3 minutos (aprox. 1 h 53 min). El tiempo total lineal (suma de todos los componentes) fue de 147.6 minutos en promedio, con una mediana de 140.4 minutos. El componente que más tiempo consumió fue el bloque de cuestionarios (100 min), seguido por los tests del responsable principal (34.2 min) y el test de vocabulario (13.3 min). La desviación estándar del tiempo total lineal fue de 39.8 minutos.

A continuación, la Tabla 2.10 detalla las estadísticas del tiempo total lineal de aplicación por región. Adicionalmente, la Tabla 2.11 muestra cómo varía la duración total lineal promedio según el número de integrantes del hogar, lo cual afecta principalmente la duración de la sección de composición del hogar en el cuestionario al responsable principal.

Tabla 2.10: Estadísticas de tiempo total lineal (en minutos) por región (datos trimados al 3%)

Región	Promedio (min)	Mediana (min)	Desv. Estándar	Mínimo (min)	Máximo (min)	N Casos
Arica y Parinacota	165,8	161,1	43,0	85,3	280,6	146
Tarapacá	138,2	130,7	37,6	82,1	279,9	189
Antofagasta	151,0	145,7	39,3	81,2	283,8	432
Atacama	143,3	138,3	39,0	81,8	279,4	165
Coquimbo	139,6	132,9	35,6	81,3	286,4	465
Valparaíso	155,1	150,5	39,0	81,4	289,6	985
Metropolitana	152,0	144,6	40,3	81,2	288,9	3.367
O Higgins	130,7	124,2	31,8	82,8	283,8	618
Maule	150,0	141,1	45,0	82,2	289,0	610
Ñuble	145,7	139,4	38,3	83,4	284,6	243
Biobío	146,9	140,3	39,2	81,4	288,8	1.036
Araucanía	136,1	127,7	40,2	81,0	278,8	555
Los Ríos	145,5	141,9	38,1	82,3	263,0	174
Los Lagos	142,5	139,0	35,7	81,4	272,0	547
Aysén	156,6	145,1	42,1	94,7	275,6	78
Magallanes	133,4	127,0	30,4	84,7	215,1	91
Total Nacional	147,6	140,4	39,8	81,0	289,6	9.701

Tabla 2.11: Estadísticas de tiempo total lineal (en minutos) por tamaño del hogar (datos trimados al 3%)

Tamaño Hogar	Promedio (min)	Mediana (min)	Desv. Estándar	Mínimo (min)	Máximo (min)	N Casos
1	122,3	123,2	29,0	82,8	171,2	8
2	145,5	136,8	41,6	81,4	287,5	763
3	146,3	138,4	40,0	81,2	288,9	2.284
4	146,4	139,2	39,5	81,0	289,2	3.229
5	146,6	139,9	38,6	81,2	289,6	2.061
6	152,5	145,2	39,9	81,4	288,8	835
7	156,9	151,6	39,4	85,9	286,4	312
8 o más	164,9	158,3	42,6	95,4	288,5	209
Total	147,6	140,4	39,8	81,0	289,6	9.701

2.3.6 Supervisión del trabajo de campo

Con el objetivo de asegurar la calidad y el cumplimiento de los protocolos durante todo el trabajo de campo, se implementaron tres vías de supervisión. En primer lugar, se contó con una supervisión externa, adjudicada a una empresa independiente para verificar de manera autónoma el cumplimiento de los estándares del levantamiento. En segundo lugar, el Centro de Microdatos (CMD), como entidad ejecutora, implementó su propio plan de supervisión interna centrado en el control de calidad continuo del trabajo de sus equipos. Finalmente, el equipo de la División de Observatorio Social (DOS) del MDSF, en su rol de contraparte técnica, realizó un proceso de observación y acompañamiento en terreno para verificar directamente la correcta aplicación de los instrumentos y protocolos. A continuación, se detalla cada uno de estos procesos.

2.3.6.1 Supervisión Externa

A fin de asegurar la calidad de las distintas etapas del proceso de recolección de la cuarta ronda de ELPI, así como de los datos generados de manera independiente, el Ministerio de Desarrollo Social contrató mediante licitación pública un servicio de supervisión externa, el cual fue adjudicado por CADEM.

Como parte de este proceso, CADEM implementó un esquema de supervisión estructurado en cuatro ejes principales: proceso de capacitación, instalación de sedes, avance del trabajo operativo y supervisión de entrevistas. Esta última contempló tanto instancias de supervisión presencial como telefónica, con el propósito de asegurar la correcta implementación de los protocolos metodológicos y resguardar la calidad de la información recolectada.

a) Supervisión de la Capacitación

El proceso de supervisión de capacitaciones tuvo como objetivo garantizar la calidad de la formación impartida a los encuestadores en la cuarta ronda de ELPI 2024. Se supervisaron 17 capacitaciones a encuestadores en sedes regionales entre el 14 y el 30 de mayo de 2024, evaluando distintos módulos clave para la correcta aplicación de la encuesta.

Las jornadas de capacitación se realizaron en sesiones de tres días consecutivos, con una asistencia promedio de 13 encuestadores por jornada. La duración efectiva de cada día de capacitación registrada por los supervisores de CADEM fue de aproximadamente siete horas, combinando presentaciones teóricas con ejercicios prácticos.

Cada módulo de la capacitación fue evaluado en términos de claridad de exposición, manejo de consultas y cumplimiento de los contenidos mínimos establecidos. En general, todos los módulos superaron el 90% de cumplimiento de los contenidos mínimos y obtuvieron calificaciones promedio superiores a seis por parte de los supervisores de CADEM. Esto incluyó los aspectos generales de la encuesta, el uso de tablets, la aplicación de los cuestionarios al responsable principal y al adolescente, así como las evaluaciones.

En cuanto a la comunicación de los aspectos administrativos, se explicaron adecuadamente las condiciones de pago a los encuestadores, incluyendo el pago mínimo por encuesta. Sin embargo, el aspecto con menor cumplimiento fue la especificación de los pagos por movilización, alimentación y viáticos, que solo se abordó en el 71% de las capacitaciones.

En lo que respecta a la infraestructura, se evaluaron positivamente las condiciones de comodidad, iluminación y disponibilidad de equipos audiovisuales. Todas las sedes contaron con el manual de trabajo de campo y tarjeteros, mientras que el 94% tuvo acceso a internet.

b) Supervisión de Instalación de Sedes

El proceso de supervisión de la instalación de sedes tuvo como objetivo verificar que las condiciones logísticas y operativas establecidas en los términos de referencia fueran

cumplidas. Para ello, se llevaron a cabo visitas presenciales a las sedes de operación de todas las regiones entre el 16 de mayo y el 7 de junio de 2024. Algunas de las supervisiones se realizaron con posterioridad a la fecha original estipulada, 22 de mayo, dada la imposibilidad de ponerse en contacto con los coordinadores de sede o dado que las sedes aún no se encontraban operativas.

Entre los aspectos evaluados en esta etapa se incluyeron la disponibilidad de equipamiento tecnológico, la habilitación de espacios adecuados para la recepción y almacenamiento de material, así como la implementación de medidas de seguridad y acceso. Los resultados de la supervisión reflejan diferencias en el nivel de cumplimiento entre las sedes. Dos de ellas (Antofagasta y Santiago) alcanzaron un 100% de cumplimiento en todos los aspectos evaluados. Seis sedes (Copiapó, San Fernando, Chillán, Temuco, Valdivia y Puerto Montt) obtuvieron entre un 93% y un 96% de cumplimiento, mientras que otras siete (Iquique, Coquimbo, Valparaíso, Talca, Concepción, Coyhaique y Punta Arenas) registraron niveles entre un 83% y un 86%. La sede de Arica presentó el menor desempeño con un 74% de cumplimiento, principalmente debido a la falta de material de escritorio, como impresoras operativas, resmas de hojas y tinta.

En términos generales, los aspectos mejor evaluados, con un 100% de cumplimiento, incluyeron la presencia de puertas y ventanas en buen estado, escritorios y sillas para los coordinadores de sede, computadores o notebooks asignados, credenciales y material didáctico. En contraste, los aspectos con menor cumplimiento (69%) estuvieron relacionados con elementos de seguridad, como muebles con chapa de seguridad para dispositivos móviles, cámaras de vigilancia y accesos restringidos mediante porterías o conserjerías. También se detectó que varios coordinadores de sede utilizaron su telefonía móvil personal en ausencia de dispositivos provistos por la organización.

c) Supervisión de avance del trabajo operativo

El proceso de supervisión del avance del trabajo operativo de la cuarta ronda de ELPI 2024 se llevó a cabo entre el 20 y el 23 de agosto, con visitas a seis sedes: Valparaíso, San Fernando, Talca, Temuco, Santiago y Concepción. La supervisión se realizó mediante entrevistas semiestructuradas con los coordinadores de sede, recopilando información clave sobre la asignación de encuestas, las tasas de respuesta y rechazo, la composición de los equipos, la logística de visitas y los factores que han afectado el desarrollo de la encuesta.

Durante las entrevistas, se consultó sobre los niveles de avance del trabajo operativo, incluyendo las tasas de logro y rechazo, así como la disponibilidad de coordinadores de grupo y encuestadores activos. En estos aspectos, no se reportaron dificultades significativas.

Al ser consultados sobre los elementos que han afectado negativamente el desarrollo de la encuesta, el problema más recurrente fue la dificultad para contactar a los encuestados, reportado como frecuente en la mitad de las sedes supervisadas. También se identificaron inconvenientes en la asignación de códigos de disposición final y, en algunos casos, en la disponibilidad de dispositivos electrónicos. Además, las condiciones climáticas adversas,

como temporales en la Región Metropolitana, Biobío y La Araucanía, interrumpieron la programación de visitas y afectaron la movilidad de los encuestadores.

En términos de gestión operativa, la supervisión permitió evaluar el uso de herramientas de seguimiento proporcionadas por el Centro de Microdatos. Todas las sedes afirmaron contar con estas herramientas y utilizarlas activamente para monitorear el estado de la muestra y coordinar el trabajo de campo.

La evaluación del cumplimiento de la tasa de logro promedió una calificación de 6,7, con sedes como Valparaíso, O'Higgins, Maule y Biobío destacando con la máxima puntuación (7). La calidad de los datos de contacto fue el aspecto con la menor calificación, con un promedio de 5,3, evidenciando problemas con direcciones desactualizadas, teléfonos incorrectos y falta de precisión en las ubicaciones rurales. Estos hallazgos coinciden con el reporte de dificultades, que identifica el contacto con los adolescentes de la muestra como el principal desafío en la recolección de datos.

d) Supervisión de Entrevistas

La supervisión de entrevistas realizada por CADEM durante la cuarta ronda de la ELPI 2024 tuvo como objetivo asegurar la calidad y cumplimiento de los protocolos metodológicos durante la aplicación del instrumento en terreno. Este proceso se implementó tanto en modalidad presencial como telefónica, concretando un total de 3.032 supervisiones, de las cuales 903 fueron presenciales (100,3% de la muestra objetivo de la supervisión presencial) y 2.129 telefónicas (59,1% de la muestra objetivo de la supervisión telefónica). Dentro de estas, se efectuaron 879 re-entrevistas de supervisión presenciales y 1.997 telefónicas .

A continuación, la Tabla 2.12 y Tabla 2.13 presentan los niveles de logro por región para cada una de las modalidades.

Tabla 2.12: Distribución de logros Modalidad Presencial

Región/área	Muestra objetivo	Logro supervisión	Tasa logro supervisión	Logro re-entrevista	Tasa logro re-entrevista
Nacional	900	903	100,3%	879	97,7%
Arica y Parinacota	15	19	126,7%	18	120,0%
Tarapacá	21	17	81,0%	17	81,0%
Antofagasta	42	15	35,7%	14	33,3%
Atacama	18	12	66,7%	12	66,7%
Coquimbo	45	45	100,0%	43	95,6%
Valparaíso	96	109	113,5%	105	109,4%
Metropolitana	283	332	117,3%	326	115,2%
O' Higgins	56	50	89,3%	47	83,9%
Maule	61	63	103,3%	61	100,0%
Ñuble	28	36	128,6%	36	128,6%
Biobío	92	112	121,7%	111	120,7%
Araucanía	52	42	80,8%	40	76,9%
Los Ríos	17	7	41,2%	7	41,2%
Los Lagos	57	29	50,9%	27	47,4%
Aysén	8	3	37,5%	3	37,5%
Magallanes	9	12	133,3%	12	133,3%

A nivel regional, se observaron variaciones importantes en las tasas de logro. En la modalidad presencial, la mayoría de las regiones alcanzaron tasas cercanas o superiores al 100%, destacando regiones como Arica y Parinacota (126,7%), Ñuble (128,6%) y Magallanes (133,3%). Sin embargo, regiones como Antofagasta (35,7%), Los Ríos (41,2%), Los Lagos (50,9%) y Aysén (37,5%) registraron un desempeño significativamente inferior debido principalmente a desafíos geográficos, problemas logísticos y dificultad para contactar a las personas responsables.

Tabla 2.13: Distribución de logros para modalidad telefónica

Región/área	Muestra objetivo	Logro supervisión	Tasa logro supervisión	Logro re-entrevista	Tasa logro re-entrevista
Nacional	3.600	2.129	59,1%	1.997	55,5%
Arica y Parinacota	59	30	50,8%	28	47,5%
Tarapacá	83	27	32,5%	23	27,7%
Antofagasta	171	77	45,0%	64	37,4%
Atacama	71	47	66,2%	44	62,0%
Coquimbo	181	84	46,4%	77	42,5%
Valparaíso	387	197	50,9%	186	48,1%
Metropolitana	1.131	785	69,4%	740	65,4%
O' Higgins	223	119	53,4%	116	52,0%
Maule	243	155	63,8%	148	60,9%
Ñuble	114	61	53,5%	58	50,9%
Biobío	368	211	57,3%	195	53,0%
Araucanía	207	129	62,3%	122	58,9%
Los Ríos	69	38	55,1%	36	52,2%
Los Lagos	227	132	58,1%	127	55,9%
Aysén	32	17	53,1%	15	46,9%
Magallanes	34	20	58,8%	18	52,9%

En la modalidad telefónica, la tasa de supervisión nacional fue del 59,1%, siendo la Región Metropolitana la que presentó el mejor desempeño (69,4%). Las regiones con desempeño más bajo en esta modalidad fueron Tarapacá (32,5%) y Antofagasta (45,0%), donde las dificultades más recurrentes estuvieron relacionadas con problemas de contacto telefónico y la disponibilidad de datos actualizados.

La duración promedio de las re-entrevistas fue de 8,3 minutos en modalidad presencial y 6,6 minutos en telefónica, pese a algunas excepciones puntuales con entrevistas considerablemente más extensas en terreno.

Respecto de los resultados de la supervisión, uno de los aspectos centrales fue la identificación de hallazgos o alertas sobre la correcta aplicación de la encuesta original. Estas alertas, definidas previamente entre el MDSF y CADEM, fueron reportadas semanalmente para su análisis. La Tabla 2.14 resume los tipos de hallazgos detectados en ambas modalidades de supervisión.

Tabla 2.14: Hallazgos de supervisión

Tipo de Hallazgo / alerta	Presencial (N)	Telefónica (N)	Total (N)
DFC alerta 113. Supervisada. No existe dirección. Código DFC 321-8	19	0	19
DFC alerta 114. Supervisada. Dirección no corresponde a un hogar. Código DFC: 451, 452, 453, 454, 462	2	0	2
DFC alerta 115. Supervisada. Nadie contestó la encuesta	2	2	4
DFC alerta 116. Supervisada. Nombre del adolescente no coincide	0	1	1
DFC alerta 117. Supervisada. No se aplicaron todos los instrumentos	11	40	51

DFC alerta 118. Entrevistado no es responsable principal del adolescente	0	1	1
DFC alerta 119. Adolescente no vivía en la dirección	1	128	129
Total Alertas	35	172	207

Como se observa en la tabla, el hallazgo más frecuente fue la alerta 119, indicando que el adolescente ya no vivía en la dirección registrada, especialmente detectado en la supervisión telefónica. También destaca la alerta 117, levantada a partir de la re-entrevista, señalando casos donde no se aplicaron todos los instrumentos correspondientes. Los casos que generaron alertas fueron reportados a el Centro de Microdatos para su revisión, además de ser sometidos a revisión en contraste con otras fuentes según lo indicado en la sección de 3.1.2.1 Validaciones Posteriores (Post-Campo).

Se realizó una comparación entre las respuestas obtenidas en la entrevista original de ELPI y las recogidas en la re-entrevista de supervisión para un conjunto de variables clave, con el fin de evaluar la consistencia de la información. La Tabla 2.15 presenta los porcentajes de coincidencia para cada variable analizada en el total de re-entrevistas realizadas.

Tabla 2.15: Coincidencia a nivel de variable por personas Presencial y Telefónica

Variable	% de consistencia	N
N° de personas que viven en la vivienda	86,4%	2.876
Responsable principal: Asiste a establecimiento educación	96,3%	2.864
Responsable principal: Nivel educacional	73,6%	2.864
Sexo Responsable Principal	97,3%	2.876
Adolescente: Edad	56,9%	2.876
Adolescente: Asiste a Establecimiento Educación	95,3%	2.875
Adolescente: Nivel educacional	78,3%	2.876
Madre Biológica Vive en el Hogar	98,7%	2.876
Padre Biológico Vive en el Hogar	95,1%	2.876

Los resultados muestran altos niveles de consistencia en la mayoría de las variables factuales, como la presencia de la madre biológica en el hogar (98,7%), el sexo del responsable principal (97,3%) o si este o el adolescente asisten a un establecimiento educacional (96,3% y 95,3% respectivamente). La variable sobre la presencia del padre biológico en el hogar también presenta una alta consistencia (95,1%).

Se observan niveles de consistencia algo menores, aunque aún mayoritarios, en variables como el número de personas que viven en la vivienda (86,4%) y el nivel educacional alcanzado por el adolescente (78,3%). El nivel educacional del responsable principal presentó una consistencia del 73,6%. Si bien esto muestra un nivel de discrepancia intermedio, este se debe mayormente a diferencias dentro de un nivel educativo, por ejemplo, media científico humanista en lugar de técnico profesional. La variable con menor coincidencia fue la edad del adolescente (56,9%), aunque el informe de supervisión externa aclara que esta se calculó a partir de la fecha de nacimiento consultada en la re-entrevista, y que las discrepancias se limitaron mayoritariamente a diferencias de un año. Estas diferencias menores pueden atribuirse a la distinta fuente de información o a cambios ocurridos entre ambas mediciones. En general, los resultados sugieren una buena calidad y consistencia de la información recolectada en la encuesta original.

Finalmente, durante la re-entrevista, se consultó a los participantes sobre su percepción del desempeño del encuestador original de ELPI. Los resultados reflejan una evaluación muy positiva. En cuanto al cumplimiento de protocolos clave, el 99,1% de los re-entrevistados confirmó haber recibido un consentimiento informado y el 99,9% indicó que se aplicaron cuestionarios tanto al responsable principal como al adolescente. Un 98,3% recordó que algunos cuestionarios fueron autoaplicados. El aspecto con un cumplimiento levemente menor fue la explicación de los objetivos de la ELPI, reportada por el 96,7% de los participantes.

La valoración subjetiva del encuestador también fue muy alta. Respecto a la claridad de la información entregada, el 93,8% de los participantes calificó al encuestador con nota máxima (7), y un 5,3% adicional con nota 6. En cuanto al trato recibido, la evaluación fue aún mejor, con un 97% otorgando nota 7 y un 2,4% nota 6. Las calificaciones inferiores a 6 fueron marginales en ambos aspectos, representando menos del 1% del total.

2.3.6.2 Supervisión CMD

De forma complementaria a la supervisión externa, el Centro de Microdatos (CMD) de la Universidad de Chile, como entidad ejecutora del levantamiento, implementó su propio Plan de Supervisión Interna. Este plan tuvo como objetivo garantizar la calidad de la recolección y minimizar posibles sesgos, evaluando continuamente el trabajo de los encuestadores. Se estructuró en dos líneas paralelas y complementarias, apoyadas por un sistema web de gestión: una supervisión de escritorio continua al 100% de las encuestas completadas y una supervisión en terreno mediante re-entrevistas a 3.008 casos, correspondientes al 30,1% de las encuestas logradas.

a) Supervisión de Escritorio

Esta línea de supervisión implicó que los coordinadores de grupo revisaran la totalidad de las encuestas marcadas como completas (CDF 110), es decir, aquellas donde se aplicaron todos los cuestionarios e instrumentos tanto al responsable principal como al adolescente. Utilizando perfiles específicos en la plataforma Survey Solutions y apoyándose en pautas y manuales, los coordinadores evaluaron la consistencia interna de las respuestas, el correcto llenado de los cuestionarios y la adecuada aplicación de las evaluaciones.

Este proceso permitió identificar y corregir errores de registro de manera oportuna. Los principales hallazgos se relacionaron con el registro en preguntas abiertas, como inconsistencias en el nombre del responsable principal entre formularios, especificaciones poco claras en preguntas de opción “otra”, falta de consistencia lógica entre preguntas relacionadas (ej. tuición compartida y residencia habitual), registro erróneo de la dependencia administrativa de establecimientos educacionales, nombres o direcciones incompletas o abreviadas de colegios (dificultando la codificación posterior), uso incorrecto del formato de hora, y confusión en conceptos como el tipo de conexión a internet o descripciones vagas de ocupaciones laborales. Ante la detección de estos problemas, los coordinadores rechazaban la encuesta en la plataforma para que el encuestador la

corrigiera, lo que a veces requería un nuevo contacto con el hogar. Una vez subsanadas las observaciones, el coordinador aprobaba la encuesta. Esta revisión se realizó de forma continua durante todo el trabajo de campo y hasta dos semanas después de su finalización, en paralelo con el proceso de codificación.

b) Supervisión en Terreno (Re-entrevistas)

Para verificar en terreno la veracidad de la información recopilada y el cumplimiento de los protocolos de aplicación, se implementó una supervisión mediante re-entrevistas a una submuestra de las encuestas completas. Aunque inicialmente se planificó controlar al 10% de las encuestas completadas por región, esta meta se redefinió, aumentando al 20% como base, con la flexibilidad de incrementar significativamente este porcentaje (hasta el 100%) para encuestadores específicos en caso de detectarse indicios o confirmarse irregularidades. Aproximadamente el 40% de estos controles se realizaron telefónicamente para complementar la cobertura.

Esta tarea fue realizada por un equipo dedicado de 42 controladores, coordinados desde cada sede y a nivel nacional. Se utilizó un cuestionario reducido, acordado con el MDSF y programado en Survey Solutions, enfocado en verificar la correcta aplicación de los instrumentos a las personas correspondientes y el cumplimiento de protocolos clave. Se estableció un protocolo de contacto estricto para los controladores, incluyendo el uso visible de credenciales, registro en la web del CMD para verificación de identidad, uso de un texto estandarizado para explicar el motivo de la visita y aplicación del cuestionario en tablet.

En total, se realizaron 3.008 re-entrevistas de supervisión interna, lo que representa un 30,1% de las 10.003 encuestas completadas a nivel nacional. Este porcentaje superior al 20% base se explica por el aumento de un 20% a un 30% de supervisión en los casos de encuestadores en los que se encontraron indicios de falsificación, y del 100% en los casos en que se encontró algún tipo de falsificación. La cobertura regional y la distribución del porcentaje de supervisión por encuestador se detallan en las Tabla 2.16 y Tabla 2.17.

Tabla 2.16: Resumen de encuestas controladas por región (Supervisión CMD)

Región	Total encuestas supervisadas	Total encuestas completas	Supervisión (%)
Arica y Parinacota	35	151	23,2%
Tarapacá	47	199	23,6%
Antofagasta	143	448	31,9%
Atacama	47	171	27,5%
Coquimbo	144	468	30,8%
Valparaíso	282	1.006	28,0%
Metropolitana	1.076	3.456	31,1%
O'Higgins	176	632	27,8%
Maule	221	651	33,9%
Ñuble	84	254	33,1%
Biobío	297	1.069	27,8%
Araucanía	172	586	29,4%
Los Ríos	64	179	35,8%
Los Lagos	178	564	31,6%
Aysén	24	78	30,8%

Magallanes	18	91	19,8%
Total	3.008	10.003	30,1%

Tabla 2.17: Distribución de encuestadores según tramo de supervisión (Supervisión CMD)

Porcentaje de supervisión	Frecuencia
Entre 10% y menos de 20%	17
Entre 20% y menos de 30%	165
Entre 30% y menos de 40%	106
Entre 40% y menos de 50%	24
Entre 50% y menos de 60%	11
Entre 60% y menos de 70%	10
Entre 70% y menos de 80%	3
Entre 90% y 100%	12
Total encuestadores	348

Como parte del análisis de la supervisión en terreno, se compararon las respuestas de 16 variables clave entre la encuesta original y la re-entrevista. Los resultados mostraron una alta consistencia general: el 59,7% de las re-entrevistas no presentó ninguna diferencia respecto a estas variables, y el promedio de discrepancias por encuesta fue bajo (0,62 preguntas), con una mediana de cero. Las diferencias menores detectadas (entre 1 y 3 inconsistencias) se revisaron caso a caso con los encuestadores y coordinadores, permitiendo rectificar información y retroalimentar al equipo; estas encuestas fueron “Aprobadas con observaciones”. El detalle de inconsistencias puede revisarse en la Tabla 2.18.

Tabla 2.18: Distribución de consistencia entre encuesta original y re-entrevista (Supervisión CMD)

Número de Diferencias	Frecuencia	Porcentaje
0	7.575	59,7%
1	3.444	27,1%
2	1.064	8,4%
3	391	3,1%
4	90	0,7%
5	78	0,6%
6	38	0,3%
7	6	0,0%
8	6	0,0%

No obstante, este proceso también fue crucial para detectar irregularidades graves, las cuales se relacionan particularmente con la aplicación incorrecta del protocolo de entrevista: aplicar cuestionarios a personas que no cumplían la definición de Responsable Principal o que el responsable contestara secciones que debían ser autoaplicadas por el adolescente. Se identificaron 5 encuestadores que incurrieron en faltas de este tipo, en las regiones de Coquimbo, Valparaíso, Maule y Metropolitana. Ante la confirmación de estas faltas, se activaron protocolos que incluyeron la revisión exhaustiva (hasta el 100%) del trabajo del encuestador involucrado, su suspensión y eventual desvinculación, y la eliminación de las encuestas afectadas (intentando recuperarlas posteriormente con otro encuestador).

Estas situaciones explican los casos de encuestadores con porcentajes de supervisión muy elevados en la Tabla 2.17. Adicionalmente, se monitorearon y abordaron casos de encuestas con tiempos de aplicación inusualmente bajos, aumentando también el porcentaje de supervisión y realizando acompañamiento en terreno. Los hallazgos y medidas adoptadas fueron reportados periódicamente al MDSF.

2.3.6.3 Supervisión MDSF

Para asegurar la calidad de los datos recopilados, el equipo de la División de Observatorio Social (DOS) del MDSF llevó a cabo un proceso de observación en terreno. Este operativo complementó las supervisiones externa y del CMD, enfocándose en la observación directa de la aplicación de los instrumentos y en recoger la percepción de los equipos regionales sobre el desarrollo del levantamiento.

El proceso se realizó de manera simultánea los días 17, 18 y 19 de julio de 2024 en seis regiones estratégicas (Arica y Parinacota, Antofagasta, Valparaíso, Metropolitana, O'Higgins y Biobío). Durante estas jornadas, se realizaron 7 entrevistas en profundidad con los coordinadores(as) regionales y se acompañó a los equipos de terreno en un total de 29 visitas a domicilio, cuyos resultados se detallan en la siguiente tabla.

Tabla 2.19: Resultados de las visitas a domicilio observadas por equipo MDSF

Región	Dirección no localizada	Entrevista aplicada	Rechazo	Total visitas
Antofagasta	4	4		8
Arica y Parinacota	3	3		6
Biobío		2		2
Metropolitana		3		3
O'Higgins		3		3
Valparaíso	3	3	1	7
Total general	10	18	1	29

A partir de este operativo de observación, se levantaron los siguientes hallazgos principales:

Hallazgos sobre la aplicación de instrumentos:

- **Desempeño general:** Se constató un alto nivel de profesionalismo y habilidades de gestión por parte de los y las encuestadoras, con un buen apego a los protocolos de entrevista y un trato cordial y respetuoso hacia los participantes.
- **Aplicación de cuestionarios:** En general, la aplicación fue fluida y correcta. Se observaron desviaciones menores, como el parafraseo ocasional de preguntas o la omisión de textos introductorios, las cuales no comprometieron la calidad de la información registrada.
- **Manejo de instrumentos:** Se verificó un uso adecuado de los materiales de apoyo, como los tarjeteros y el cuadernillo del test TVIP.

Hallazgos sobre la gestión del terreno a partir de entrevistas con coordinadoras:

- Principales facilitadores: Los factores que más ayudaron al levantamiento fueron la buena disposición de las familias a participar (muchas recordaban rondas anteriores), la entrega de incentivos (giftcard y obsequios), y la motivación de los equipos de terreno por la relevancia de la encuesta.
- Principal obstáculo: De manera transversal, el mayor desafío reportado fue la dificultad para establecer el contacto inicial con los hogares, debido a la desactualización de direcciones y, especialmente, de los números telefónicos del marco de contacto. Los equipos de terreno debieron desplegar un alto nivel de gestión para localizar a los participantes, utilizando diversas estrategias complementarias.
- Otros desafíos: También se mencionaron la larga duración de la entrevista, que en ocasiones requería agendar una segunda visita, y la complejidad logística de transportar la gran cantidad de materiales de campo.

En síntesis, la supervisión del MDSF confirmó la correcta ejecución de los protocolos en terreno y validó que la principal dificultad del levantamiento se centró en la fase de localización de la muestra, más que en la obtención de la cooperación una vez establecido el contacto.

Capítulo 3

PROCESAMIENTO

3.1 REVISIÓN Y VALIDACIÓN DE DATOS

El proceso de revisión y validación de los datos de ELPI 2024 fue una tarea continua y colaborativa, realizada tanto por el Centro de Microdatos (CMD) como por el equipo de la División de Observatorio Social (DOS), durante y después del período de levantamiento. El objetivo fue asegurar la calidad, consistencia e integridad de la información recopilada.

3.1.1 Validación de datos por el Centro de Microdatos

El CMD implementó un sistema de validación concurrente al trabajo de campo. A medida que las encuestas eran completadas y sincronizadas, pasaban por un proceso de revisión que incluía:

1. **Revisión de Escritorio (Coordinadores de Grupo):** Como se detalló en la sección de Supervisión CMD, los coordinadores revisaron el 100% de las encuestas completas (CDF 110) en la plataforma Survey Solutions, verificando consistencia interna básica y correcto llenado, rechazando encuestas para corrección por parte del encuestador si era necesario.
2. **Aplicación de Mallas de Validación Programáticas:** Se aplicaron rutinas automatizadas para detectar inconsistencias lógicas y errores específicos definidos en una malla de validación, tanto programadas en el software de aplicación como revisiones ex post.
3. **Recuperación de Información:** Cuando se detectaban errores o inconsistencias que no podían resolverse internamente, se activaba un protocolo para que el equipo de terreno contactara nuevamente al hogar entrevistado y corrigiera la información.

La Tabla [3.1](#) resume el estado final de las validaciones realizadas por CMD sobre los errores detectados en general. Cerca de la mitad de los casos validados requirieron algún

tipo de corrección, incluyendo los casos en que se realizó recuperación de información desde terreno.

Tabla 3.1: Estado final de validación de errores generales (CMD)

Estado	Cantidad	%
Validado sin corrección	3.002	52,9%
Validado con corrección	2.672	47,1%
Total	5.674	100,0%

Centrándose específicamente en los errores detectados por la malla de validación programática, la distribución de correcciones fue similar, como muestra la Tabla 3.2.

Tabla 3.2: Estado final de validación para errores de malla programática (CMD)

Estado	Cantidad	%
Validado sin corrección	2.123	54,6%
Validado con corrección	1.763	45,4%
Total	3.886	100,0%

La Tabla 3.3 detalla la frecuencia de los errores específicos más prevalentes encontrados por la malla de validación del CMD. Los errores más comunes se relacionaron con la consistencia de los ingresos por pensión de alimentos (y5_e2), rangos plausibles para las horas de sueño del adolescente (tot_hrs_sueno...), valores bajos de ingreso laboral (y1_e1), y horarios atípicos de despertar o dormir (pc1...). La detección y corrección de estos errores fue fundamental para mejorar la calidad de los datos.

Tabla 3.3: Frecuencia de errores específicos en malla de validación (CMD)

Error	Descripción	N	Porcentaje
y5_e2	Si en p7/m7 responde sí, y en p8/m8 ese acuerdo se cumple, y en p9/m9 responde que sí, debiera declarar ingreso en esta pregunta en receptor (17 o menos el receptor es el RP, 18 y más el receptor es el joven); osea no puede tener cero ingreso. Antes declaró recibir pensión de alimentos o aporte económico, revise respuesta ingresada	689	17.73%
tot_hrs_sueno_fin_p	Total horas de sueño en día de fin de semana muy alto	434	11.17%
y1_e1	Revisar valores bajos (103.625 para trabajadores dependientes y 42.583 para independientes). Revise el monto ingresado	328	8.44%
tot_hrs_sueno_fin_p	Total horas de sueño en día de fin de semana muy bajo	263	6.77%
pc1_c	Hora a la que se despierta en día de fin de semana es muy temprano	259	6.66%
tot_hrs_sueno_sem_p	Total horas de sueño en día de semana muy bajo	252	6.48%
tot_hrs_sueno_sem_p	Total horas de sueño en día de semana muy alto	234	6.02%
pc1_c	Hora a la que se despierta en día de fin de semana es muy tarde	220	5.66%
pc1_b	Horario atípico para dormir en día de semana	213	5.48%
pc1_a	Hora a la que se despierta en día de semana es muy temprano	202	5.20%
pc1_a	Hora a la que se despierta en día de semana es muy tarde	178	4.58%
pc1_d	Horario atípico para dormir en día de fin de semana	171	4.40%
h5_e2	Parejas declaradas con distinto estado civil. Hay parejas con distinto estado civil, revisar.	54	1.39%
Otros	Otros Errores	389	10.00%
Total	Total Errores	3886	

Finalmente, como parte del proceso de preparación de las bases de datos finales, CMD realizó una edición programática específica para la variable g7 ("¿Cuántos años tenías

cuando tuviste tu primera relación sexual voluntaria?”). Dado que el cuestionario no incluía explícitamente una opción “No responde”, los valores registrados como “0” en esta variable fueron recodificados al código estándar de valor perdido “-99” para indicar la ausencia de respuesta válida.

3.1.2 Validación de datos por el Ministerio de Desarrollo Social

Paralelamente al trabajo del CMD, el equipo de la DOS del MDSF implementó su propio proceso de validación, con el objetivo inicial de contar con una programación independiente que permitiera contrastar resultados y maximizar la detección de posibles errores. Esta validación se basó en la construcción de una “malla de validación” utilizando el paquete `validate` de R, la cual se aplicó semanalmente a los datos recibidos del CMD.

Esta malla definió reglas lógicas que los datos debían cumplir en tres áreas principales:

1. Reglas de Universo: Verificaban que las preguntas fueran respondidas por quienes correspondía según los flujos del cuestionario y las condiciones habilitantes.
 - Para preguntas sin condición habilitante (ej. edad), se chequeaba la ausencia de datos (`is.na(edad)`).
 - Para preguntas con condición habilitante, se creaban dos reglas: una para detectar casos que debían responder y no lo hicieron (ej. `is.na(pc8__1) & pc7 == 1`), y otra para detectar casos que respondieron sin cumplir la condición (ej. `!is.na(pc8__1) & pc7 != 1`).
2. Reglas de Rango: Aseguraban que los valores de las variables estuvieran dentro de los límites lógicamente posibles, sin evaluar aún la plausibilidad o la presencia de outliers (que se tratarían como errores de contenido o alertas).
 - Categóricas: Se verificaba que el valor perteneciera al conjunto de códigos válidos (ej. `!sexo %in% c(1, 2)`).
 - Continuas: Se comprobaba que el valor estuviera dentro de un intervalo lógico (ej. `!(ingreso_trabajo >= 0 & ingreso_trabajo <= 99999999)`).
 - Fechas: Se validaba que estuvieran dentro de un período temporal razonable (ej. `!(fecha_nacimiento >= '2006-01-01' & fecha_nacimiento <= '2009-09-01')`).
 - Texto (Carácter): Se evaluaba que las respuestas abiertas tuvieran una longitud mínima de caracteres razonable (ej. `stringr::stri_length(ocupacion_especifique) < 3`).
3. Reglas de Contenido (Consistencia): Evaluaban la coherencia lógica entre distintas variables dentro de una misma encuesta. Estas reglas fueron desarrolladas de forma colaborativa entre MDSF y CMD. Ejemplos incluyen verificar la consistencia entre la edad y el curso escolar, la relación de parentesco y las edades de los miembros del hogar (por ejemplo, que no hubiera menos de 11 años de diferencia entre padres e hijos), o la coherencia entre preguntas sobre tuición y residencia habitual.

Los resultados de la aplicación semanal de esta malla de validación por parte de la DOS generaban reportes de errores que eran enviados al CMD. Estos reportes servían como insumo para que el CMD verificara los casos señalados, realizara las correcciones pertinentes en terreno (si era necesario y posible) y/o ajustara sus propios procesos de validación y limpieza.

3.1.2.1 Validaciones Posteriores (Post-Campo)

Una vez finalizado el trabajo de campo y recibida la base de datos preliminar por parte del CMD, la DOS realizó validaciones adicionales utilizando fuentes externas y datos históricos:

- **Comparación con Registros Administrativos:** Se verificó la consistencia de datos clave como el nombre y la fecha de nacimiento del adolescente con registros oficiales disponibles, para confirmar la identidad del participante.
- **Cruce con Base ELPI 2010:** Se comparó la información de la ronda actual con la registrada en la primera ronda (2010), chequeando principalmente:
 - La consistencia de la edad del adolescente (considerando el tiempo transcurrido).
 - La diferencia de edad calculada en 2024 entre el adolescente y sus padres/madres versus la diferencia registrada en 2010.

Estas validaciones posteriores, que además se compararon con las alertas de supervisión externa, no generaron alertas significativas que indicaran problemas sistemáticos de identificación de los adolescentes entrevistados o indicios de falsificación de encuestas. Las discrepancias encontradas fueron mínimas y atribuibles a errores puntuales de registro o a la dinámica propia de los datos a lo largo del tiempo. En particular, 452 casos presentaron una sola alerta, la mayor parte relacionadas con inconsistencias entre rondas en el registro de la edad de los padres. Solo 19 casos presentaron dos discrepancias, los cuales fueron revisados manualmente.

Este proceso de revisión y validación, combinando chequeos automáticos, revisión manual, recuperación en terreno y cruces con datos externos, fue esencial para asegurar la calidad y confiabilidad de la base de datos final de ELPI 2024.

3.2 CLASIFICACIÓN Y CODIFICACIÓN

El procesamiento de los datos recopilados incluyó una etapa de clasificación y codificación, orientada a estandarizar las respuestas abiertas y facilitar el análisis posterior. Este proceso fue llevado a cabo inicialmente por el Centro de Microdatos (CMD) y complementado con validaciones y revisiones adicionales por parte del equipo de la División de Observatorio Social (DOS).

El equipo del CMD, compuesto por cuatro codificadores, se encargó de la codificación de las preguntas abiertas principales durante y después del trabajo de campo. Se realizó una capacitación inicial para refrescar y consensuar los criterios de codificación, utilizando los manuales y clasificadores correspondientes, junto con ejemplos prácticos. La codificación se realizó de manera iterativa a medida que las encuestas eran completadas en terreno.

Las principales variables codificadas por el CMD fueron:

- Rama de Actividad Económica del Responsable Principal: Pregunta o8a “¿A qué se dedica o qué hace el negocio, empresa o institución donde %persona% trabaja?” se codificó usando el Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES).
- Ocupación del Responsable Principal: Pregunta o8b “¿Qué hace %persona% en su trabajo o negocio principal?” se codificó utilizando el Clasificador Chileno de Ocupaciones (CIUO 08.CL).
- Expectativas Laborales del Adolescente: Preguntas cn7 (según el responsable principal: “¿Qué trabajo Ud. piensa que realizará %adolescente% cuando termine sus estudios?”) y e8 (según el adolescente: “¿Qué trabajo piensas que harás cuando termines tus estudios?”).
- Establecimiento Educativo: Pregunta en4 (para identificar el colegio o establecimiento de educación superior). Se utilizó el Rol Base de Datos (RBD) de cada establecimiento para su codificación.

El equipo codificador revisó las respuestas para asegurar que contuvieran información suficiente y gestionó la recuperación de información con los encuestadores cuando fue necesario. La Tabla 3.4 resume los resultados cuantitativos de este proceso inicial de codificación realizado por el CMD.

Tabla 3.4: Resumen de codificaciones realizadas por CMD

	Rama de Actividad RP (o8a)	Ocupación RP (o8b)	Expectativa Laboral RP (cn7)	Establecimiento Ed. (en4)	Expectativa Laboral Ado (e8)
Codificaciones conseguidas	7.020	7.041	9.373	9.567	9.054
NS / NR / Sin código	21	0	604	0	942
Codificaciones totales	7.041	7.041	9.977	9.567	9.996

Posteriormente, el equipo de la DOS realizó procesos adicionales de validación y refinamiento sobre las codificaciones:

1. Validación de Códigos CAENES y CIUO: Las codificaciones de ocupación y actividad económica (o8a, o8b) realizadas por el CMD fueron contrastadas con los resultados de modelos de codificación automática desarrollados por el Instituto Nacional de Estadísticas (INE) con base en la encuesta CASEN. Los casos donde se detectaron discrepancias entre la codificación manual y la automática fueron revisados individualmente para asegurar la asignación del código más preciso.

2. **Revisión de Categorías Residuales (“Otra”)**: Se llevó a cabo una revisión exhaustiva de las respuestas registradas en la opción “Otra (especifique)” para un conjunto seleccionado de preguntas cerradas. El objetivo fue determinar si estas respuestas podían ser reclasificadas en alguna de las categorías preexistentes, caso en el cual fueron reasignadas. La cantidad de casos que fueron recodificados a partir de este proceso puede verse en la tabla Tabla 3.5.

Las variables sometidas a este proceso fueron:

- p3: Razón principal por la que el padre no vive con el adolescente.
- m3: Razón principal por la que la madre no vive con el adolescente.
- e2: Razón principal por la que el adolescente no asiste a un establecimiento educacional.
- home9: Otras acciones disciplinarias ante una falta de respeto del adolescente.

Adicionalmente, como parte de esta revisión, se identificaron casos donde la razón por la que el padre o la madre no vivía en el hogar era su **fallecimiento**. Dado que la respuesta original era “Otra”, se habían aplicado incorrectamente las preguntas subsiguientes sobre contacto, pensión alimenticia y régimen de visitas. Por ello, en los casos recodificados (2 para padres y 1 para madres), se procedió a editar el flujo de la encuesta, asignando un valor perdido (NA) a todas las preguntas que no correspondían aplicarse.

Tabla 3.5: Casos recodificados desde la categoría ‘Otra’ a categorías preexistentes

Variable	Descripción	Casos Recodificados
p3	Razón principal por la que el padre no vive con el adolescente	119
m3	Razón principal por la que la madre no vive con el adolescente	32
e2	Razón principal por la que el adolescente no asiste a un establecimiento educacional	14
home9	Otras acciones disciplinarias ante una falta de respeto del adolescente	203

3.3 EDICIÓN DE DATOS

Una vez finalizada la etapa de validación, se procedió a la edición de la base de datos ELPI 2024. Este proceso implicó la corrección programática de inconsistencias específicas y la aplicación de métodos estadísticos para tratar la no respuesta en variables de ingresos.

3.3.1 Edición de Variables

Tras la validación de los datos, el CMD realizó un proceso de edición programática para corregir inconsistencias específicas detectadas durante la revisión. Estas ediciones buscan mejorar la calidad y coherencia interna de la base de datos final. Las principales correcciones fueron:

- **Edad de la primera relación sexual voluntaria (g7):** El cuestionario para la pregunta g7 (“¿Cuántos años tenías cuando tuviste tu primera relación sexual voluntaria?”) no incluía una opción explícita de “No responde”. Por ello, los casos donde no se obtuvo esta información y fueron registrados con un valor de “0” se recodificaron a “-99”, el código estándar en ELPI para “No responde” o “No sabe”. Esta edición asegura un tratamiento consistente de la no respuesta para esta variable.
- **Inconsistencia entre Edad y Nivel Educativo:** Se detectaron 15 casos de adolescentes que declaraban estar cursando o haber cursado años de educación superior (universitaria o profesional) que resultaban inconsistentes con su edad. Dado que no fue posible recuperar esta información desde terreno, se optó por recodificar el curso a “-88” (No sabe), manteniendo el nivel educativo reportado para no perder información, pero señalando la inconsistencia en el detalle del avance curricular.
- **Inconsistencias en la Composición del Hogar:** Se realizaron ediciones en variables de parentesco para resolver contradicciones lógicas en la estructura del hogar:
 - En tres hogares se registraron personas con el parentesco “Hijo(a) sólo del esposo(a) o pareja”, pero no existía una pareja declarada para el jefe o jefa de hogar. En estos casos, el parentesco se corrigió a “Hijo(a) sólo del jefe(a) de hogar”.
 - En otros tres casos, se identificaron personas registradas como “Suegro(a)” del jefe de hogar, pero cuya edad era considerablemente menor a la de este y similar a la de sus hijos. Asumiendo un error de registro, el parentesco fue editado a “Yerno/Nuera”.

3.3.2 Imputación de Ingresos

La ausencia de respuesta en las preguntas sobre ingresos es un desafío habitual en las encuestas a hogares. Para abordar los datos faltantes en ELPI 2024 y permitir análisis más completos de la situación económica, se implementó un procedimiento de imputación para las principales corrientes de ingreso, siguiendo una metodología adaptada de la utilizada en la encuesta CASEN. Este proceso fue llevado a cabo por el Ministerio de Desarrollo Social.

El proceso de imputación se focalizó en dos componentes del ingreso:

1. **Ingresos del Trabajo (y1):** Correspondiente al monto mensual que las personas de 15 años o más declaran haber recibido por su ocupación, ya sea como sueldo, bonos, ganancias de un negocio propio o actividad independiente.
2. **Ingresos por Jubilaciones y Pensiones (y2):** Se refiere al monto mensual percibido por conceptos como jubilación o pensión de vejez, invalidez, orfandad, montepío o viudez.

Aunque el cuestionario ELPI indaga sobre cinco corrientes de ingreso, la imputación se concentró en estas dos debido a su importancia en la composición del ingreso total y la disponibilidad de variables auxiliares robustas para modelar la no respuesta.

El método principal empleado fue la imputación por medias de celdas (mean cell imputation). Esta técnica estadística busca asignar un valor de ingreso plausible a aquellas personas que no respondieron (receptores), utilizando la información proporcionada por personas con características similares que sí respondieron (donantes). El proceso general es el siguiente:

1. **Identificación de Casos a Imputar:** Se seleccionan las personas que no proporcionaron un monto válido (registraron “No sabe” o “No responde”).
2. **Selección de Donantes Válidos:** Se utiliza como base a las personas que sí declararon un ingreso. Para evitar la distorsión por valores extremos, se excluyen los ingresos de los donantes que se encuentran en los percentiles más bajos (1%) y más altos (99%) de la distribución.
3. **Creación de Celdas de Imputación:** Se definen grupos homogéneos de donantes y receptores basados en características sociodemográficas y laborales, como ubicación geográfica, rama de actividad, categoría ocupacional, nivel educativo, edad y sexo.
4. **Proceso Iterativo de Asignación:** La imputación se realiza de forma secuencial. Primero, se buscan donantes que coincidan exactamente con el receptor en todas las características. Si no es posible, las condiciones se relajan progresivamente en iteraciones sucesivas, agrupando categorías o eliminando variables del cruce para asegurar el “match”.
5. **Cálculo y Asignación del Ingreso:** Al receptor se le asigna el promedio del ingreso observado en el grupo de donantes (ponderado por el factor de expansión), una vez que se encuentra una celda con suficientes donantes.

Tratamiento por tipo de ingreso:

- **Ingresos del Trabajo:** Se imputó el ingreso a las 1.041 personas que, estando ocupadas, no reportaron un monto. El proceso se realizó de forma diferenciada para trabajadores dependientes e independientes.
- **Ingresos por Jubilación y Pensión:** Se adoptó un enfoque más conservador para evitar una posible sobreestimación de los ingresos del hogar. Del total de 221 personas que no respondieron a esta pregunta, el tratamiento fue diferenciado:
 - Se imputó un monto de pensión a las 119 personas que se encontraban en edad legal de jubilar (60 años o más para mujeres, 65 o más para hombres).
 - A las 102 personas restantes que no estaban en edad de jubilar, se les asignó un ingreso de cero, al considerarse que la ausencia de respuesta era más probablemente una ausencia de pensión (dada la baja prevalencia de pensiones de invalidez u orfandad en esas edades).

Como resultado de este proceso, se generaron las variables y1_corr (ingreso del trabajo corregido) e y2_corr (ingreso por jubilaciones y pensiones corregido). La siguiente tabla resume el impacto de la imputación en las estadísticas descriptivas de ambas corrientes de ingreso.

Tabla 3.6: Impacto de la imputación en los ingresos del trabajo y pensiones

Estadístico	Ingresos del Trabajo	Ingresos por Jubilación/Pensión
N° de casos con no respuesta	1.041	221
- Imputados a un valor (>0)	1.041	119
- Asignados a cero	0	102
N° de perceptores (original)	40.086	2.669
N° de perceptores (corregido)	41.127	2.788
Media (perceptores, original)	644.196	249.948
Media (perceptores, corregido)	644.808	251.164
Mediana (perceptores, original)	500.000	214.000
Mediana (perceptores, corregido)	500.000	214.000

La imputación de ingresos es un paso fundamental para mejorar la completitud y consistencia de la base de datos, permitiendo análisis más precisos sobre la situación socioeconómica de los hogares y adolescentes de la encuesta ELPI.

3.4 CÁLCULO DE NUEVAS VARIABLES

Con el objeto de facilitar el uso de las variables que pudieran ser de interés y utilidad para los usuarios, se procedió a crear o recodificar una serie de variables.

Entre dichas variables se encuentran, entre otros:

- Número de personas en el hogar
- Personas nacidas fuera de Chile
- Condición de actividad
- Tipo de hogar
- Pertenencia a pueblo indígena
- Escolaridad
- Nivel educacional
- Índice de hacinamiento
- Subsidios monetarios
- Ingreso autónomo
- Ingreso total
- Decil de ingresos
- Quintil de ingresos
- Identificación de participación en rondas

Tanto las consideraciones para la construcción de variables como su programación se puede consultar en detalle en el *Manual de Uso de base de datos*, disponible en <https://observatorio.ministeriodesarrollosocial.gob.cl/elpi-cuarta-ronda>.

3.5 CÁLCULO DE FACTORES DE EXPANSIÓN

Con el objetivo de que la muestra obtenida en ELPI resulte representativa de su población objetivo, el Instituto Nacional de Estadística construyó tres factores de expansión:

- **Factor de expansión transversal 2024:** diseñado para representar a la muestra obtenida en el año 2010. Compuesto por (1) un ponderador de selección, el que refleja las probabilidades iniciales de inclusión de cada unidad muestral en la muestra 2010, (2) ajuste por no contacto, para que unidades contactadas pudieran representar a unidades no contactadas, (3) ajuste de no respuesta, para que quienes respondieran pudieran representar a quienes no respondieron, (4) suavizamiento, para controlar valores extremos que pudieran distorsionar resultados, y (5) ponderador de calibración, que tiene por objetivo obtener estimaciones más precisas y representativas de la población en distintos dominios del estudio.
- **Factor de expansión longitudinal de panel:** diseñado con el objetivo de reflejar de mejor manera la trayectorias de niños, niñas y adolescentes que participaron en todas las olas del estudio. Se encuentra compuesto por (1) un ponderador de selección, que utilizó al factor de expansión transversal del año 2010, (2) ajuste de no respuesta, para que quienes respondieran pudieran representar a quienes no respondieron, (3) suavizamiento, para controlar valores extremos que pudieran distorsionar resultados, y (4) ponderador de calibración, que tiene por objetivo obtener estimaciones más precisas y representativas de la población en distintos dominios del estudio.
- **Factor de expansión longitudinal año 2010-2024:** diseñado para entender cómo se comportó la población que solo logró ser contactada en la primera y en la última ronda de ELPI. Se encuentra compuesto por (1) un ponderador de selección, que utilizó al factor de expansión transversal del año 2010, (2) ajuste de no respuesta, para que quienes respondieran pudieran representar a quienes no respondieron, (3) suavizamiento, para controlar valores extremos que pudieran distorsionar resultados, y (4) ponderador de calibración, que tiene por objetivo obtener estimaciones más precisas y representativas de la población en distintos dominios del estudio.

Para más detalles sobre la construcción de factores de expansión, consultar documento *Informe metodológico de construcción de factores de expansión*, disponible en <https://observatorio.ministeriodesarrollosocial.gob.cl/elpi-cuarta-ronda>.

3.6 ARCHIVOS FINALES DE DATOS

Las bases de datos de ELPI se encuentran disponibles para uso público a través de archivos descargables en diversos formatos: Stata (.dta), SPSS (.sav) y R (.RData). Junto con ello, se encuentra disponible el Libro de Códigos con la descripción de las

variables de las bases de datos, documentación relativa a la metodología de la encuesta, los cuestionarios aplicados, y el manual de uso de la base de datos. Toda la información de la encuesta puede ser descargada desde el sitio web del Observatorio Social <https://observatorio.ministeriodesarrollosocial.gob.cl/elpi-cuarta-ronda>, así como desde el Banco Integrado de Datos <https://bidat.midesof.cl/directorio/Encuestas>.

Con el objetivo de resguardar la confidencialidad de la información recolectada y evitar el acceso a información que facilite la identificación individual de las personas encuestadas, en el procesamiento de las bases se aplicaron los procesos de innominación y anonimización.

De esta manera, las bases de datos tienen el carácter de innominada e indeterminada, es decir, en las bases no se hace referencia expresa ni directa a ninguna persona natural o jurídica, y han sido sometidas a un procedimiento de eliminación de todas las variables de identificación geográfica e información personal que permitan individualizar a las personas en el momento de la entrevista. Adicionalmente, en el proceso de anonimización de datos se evaluó el riesgo de identificación de las personas con la divulgación de la información, aplicando las medidas necesarias para mitigar ese riesgo.

Para esta ronda, se publican tres bases de datos principales:

1. Base de datos del Responsable Principal: Contiene la información del cuestionario aplicado al responsable principal, incluyendo las características de todos los integrantes del hogar.
2. Base de datos del Adolescente: Incluye las respuestas del cuestionario autoaplicado por el o la adolescente.
3. Base de datos de Evaluaciones: Reúne los resultados de las escalas y tests aplicados, tanto al responsable principal como al adolescente.

Para más detalles sobre las características de las bases de datos, consultar el *Manual de Uso de base de datos*, disponible en <https://observatorio.ministeriodesarrollosocial.gob.cl/elpi-cuarta-ronda>.