

Casen en Pandemia 2020 CAPI y Aseguramiento de Calidad (QA)

Antecedentes

La recolección de datos a través de dispositivos móviles está siendo ampliamente utilizada en instituciones estadísticas para la aplicación de encuestas socioeconómicas, debido a que, de acuerdo con la evidencia encontrada, incorpora ventajas respecto a levantamientos en papel, en especial cuando se manejan grandes volúmenes de encuestas y encuestados.

El Ministerio de Desarrollo Social y Familia ha desarrollado dos estudios que han permitido evaluar el uso de dispositivos móviles en la aplicación de la Encuesta Casen. El primer estudio realizado por el Centro de Microdatos (CMD) se inició el 5 de diciembre del año 2016 y los resultados finales se entregaron en agosto 2017. El segundo estudio realizado por el Centro de Encuestas y Estudios Longitudinales (CEEL) se desarrolló dentro de las actividades del Levantamiento de la Encuesta Casen 2017, específicamente en la prueba de campo. Este estudio se inició en la segunda quincena de agosto del año 2017 y los resultados se presentaron el 19 de octubre del año 2017.

Este documento, además de los resultados de los estudios antes mencionados, incorpora el análisis de experiencias internacionales en los distintos procesos del desarrollo de una encuesta.

Finalmente, se describen las diferentes pruebas de aseguramiento de calidad o QA¹, las cuales permiten asegurar el correcto funcionamiento del instrumento en un contexto de aplicación por medio de dispositivos electrónicos. Estas pruebas se realizaron tanto por parte del contratista como por parte del Ministerio durante el proceso de diseño del cuestionario de Casen en Pandemia 2020.

¹ *Quality Assurance*, por sus siglas en inglés.

1. Evidencia internacional sobre cambio de modalidad PAPI a CAPI

Desde fines de los años ochenta, y más fuertemente desde la década de los noventa, las principales encuestas en los países desarrollados han migrado de entrevistas cara a cara utilizando cuestionarios aplicados en papel (Paper and Pencil Interview-PAPI) a entrevistas utilizando métodos electrónicos como computadores, PDAs y Tablets (Computer Assisted Personal Interviewing-CAPI). La utilización de medios electrónicos se ha acrecentado en gran medida debido a que los costos de tecnología han disminuido². Además, diversos estudios dan cuenta de la superioridad de CAPI para la recolección de datos en encuestas de carácter socioeconómico, debido a la evidente reducción de errores en los procesos de recolección de datos y procesamiento, al agregar rutinas automatizadas y verificación de validaciones.

A continuación, se presentan los principales elementos relevados en la literatura, referidos a las etapas de diseño e implementación, procesos, gestión, resultados y costos³ en un proyecto de traspaso de PAPI a CAPI, y que han sido considerados en el traspaso realizado en el marco de Casen.

1.1. Diseño e implementación

Un estudio desarrollado por Caeyers et al. (2010), señala que el éxito de CAPI depende, principalmente, del esfuerzo gastado en las etapas previas de diseño y planificación, por lo que fijar etapas y adecuados tiempos en la carta Gantt resulta de vital importancia.

En relación con la implementación, se destacan algunos elementos relevantes a tener en consideración:

- Hardware y software. En la mayoría de las pruebas realizadas⁴ se concluye que el tipo de tecnología tiene efectos, por lo que la evaluación del hardware y software⁵ empleado es fundamental. La tecnología influye en el flujo de la entrevista, reglas de validación, la transferencia de datos y el diseño de la pantalla, entre otros.
- Equipo multidisciplinario. Para el desarrollo y evaluación del software se requiere contar con un equipo multidisciplinario que incluya investigadores, personal informático, equipo de terreno y de diseño.
- Usabilidad. El instrumento CAPI necesita previamente una evaluación cuidadosa de su usabilidad a través de pruebas realizadas por un equipo multidisciplinario, o bien a través de pruebas de campo específicas que evalúen la configuración y características del software, distintas a una prueba campo tradicional, por lo que se requieren pruebas más exhaustivas que un cuestionario en papel.

² Por ejemplo, en lo que refiere a la instalación de servidores, acceso a equipos con mejores especificaciones técnicas (duración de la batería, tamaño y brillo de la pantalla, etc.), almacenamiento de la información, entre otros.

³ En el Anexo 2 se presenta análisis comparativo entre CAPI y PAPI en otros ámbitos.

⁴ Jean M. (1993), Marek et al. (2000) y Bixby et al. (2005).

⁵ Rosero et al. (2005) incluye 12 consideraciones de software para CAPI.

1.2. Procesos operativos

Entre las principales ganancias en la implementación de un proyecto en modalidad CAPI, en relación con uno en modalidad PAPI, se encuentran los procesos operativos, muchos de los cuales son automatizados, permitiendo menos errores y mayor rapidez en el procesamiento de los datos:

- Automatización de rutinas. CAPI permite el control y precisión total del flujo de cuestionarios complejos, incluyendo patrones de saltos, filtros y *rosters* (listas). El encuestador es guiado con las preguntas que corresponde aplicar y cualquier respuesta faltante o fuera de rango es advertida al instante.
- Validación de información en terreno. CAPI incluye el uso de validadores de consistencia, que alertan en forma temprana a los encuestadores de advertencias/alertas o errores. Estas verificaciones se ejecutan durante la entrevista y, por lo tanto, se pueden corregir directamente con el encuestado. Estudios⁶ confirman la mejora en la calidad de los datos con CAPI, reduciendo significativamente el número de errores.
- Administración del material de terreno. CAPI consolida en un solo dispositivo información que generalmente se encuentra separada en varios documentos, tarjetas y papeles en general. Por ejemplo, manual del encuestador, hoja de ruta, tarjeteros, imágenes, aclaraciones de preguntas, cuestionario, entre otras.
- Supervisión de encuestadores. CAPI permite hacer seguimiento a las coordenadas de terreno realizado por el encuestador, así como disponer de indicadores de terreno en forma anticipada.
- Tiempo de aplicación. En este punto la evidencia no es concluyente. Existen pruebas que concluyen que CAPI posee tiempos más bajos de aplicación (10% menos)⁷, debido a la eficiencia en el flujo del cuestionario, pero también existe evidencia que revelan que CAPI toma más tiempo que PAPI, principalmente cuando existe registro de muchos miembros en un hogar, debido a la velocidad de tipeo versus la escritura⁸. Cabe mencionar que el tiempo de aplicación depende también de la cantidad de preguntas abiertas que posee el cuestionario, dado que generalmente no existe una evaluación de las capacidades de digitación de los encuestadores.
- Percepción y confianza del encuestado. La evidencia no es concluyente. Algunos estudios no encontraron diferencias en la percepción del encuestado (grado de intimidación y percepción de confidencialidad) entre CAPI y PAPI⁹.

⁶ Caeyers et al. (2012); Martin J. (1993); Martínez M. y Moratilla T. (2000); Laurie H. (2003); Rosero L. et al. (2005); Hunter C. y Landreth A. (2005); Schraepfer J., et al. (2010); Díaz de Rada V. (2015)

⁷ Choumert et al. (2016)

⁸ Marek et al. (2000)

⁹ Martínez M. y Moratilla T. (2000)

1.3. Medidas de gestión operativa

Respecto a la tasa de respuesta, los resultados no son concluyentes. Nicholls y Martin (1997), encuentran un aumento en la tasa de respuesta en formato CAPI, fundamentalmente por reducción de errores y por ende mayor cantidad de cuestionarios completos. Laurie (2003) encuentra en general las mismas tasas en proceso de cambio a CAPI en el British Household Panel Study¹⁰, con la excepción del grupo de adultos que no habían sido entrevistados antes y que experimentan una mayor tasa de no respuesta en los ítems de ingreso.

En general, utilizando CAPI se disminuye la cantidad de datos faltantes, sin embargo, se presentan tasas de no respuesta a los ítems de ingresos y salarios levemente mayores. Una posible explicación que plantea Laurie (2003) se debe a las dificultades de ingresar montos en el dispositivo, sumado a una posible desconfianza de encuestados adultos que poseen menor interacción con la tecnología. Por otro lado, Schraepler et al. (2010) señala que una posible explicación, es que los encuestadores no confían en CAPI y transmiten sus temores sobre la tecnología.

1.4. Resultados de las estimaciones

En general no existe una evidencia empírica concluyente respecto a diferencias en resultados estadísticos entre PAPI y CAPI. Esto, principalmente, por las dificultades metodológicas y operativas que implica la comparabilidad entre ambas estrategias. Sin embargo, de una prueba experimental realizada por el Banco Mundial¹¹ es posible señalar:

- Se evaluaron las respuestas en indicadores relevantes como: actividad económica, situación conyugal, ocupación y rama económica, y no se encontró diferencia estadísticamente significativa según modo.
- Respecto al resultado de los indicadores principales del estudio, se evidencia que las medidas de consumo del hogar son más altos en modo PAPI, se demuestra mayor pobreza en modo CAPI y la diferencia es estadísticamente significativa, así como también se concluye una menor desigualdad en modo CAPI. Finalmente, la elasticidad del ingreso del consumo de alimento es menor en modo CAPI.

1.5. Costos

En relación con los costos, los estudios dan cuenta que CAPI requiere mayores costos de implementación tanto para hardware (Tablet, PDA, PC portátiles, servidores, entre otros) como para el diseño y desarrollo del software. Adicionalmente, los proyectos en esta modalidad producen un ligero aumento neto en los costos de capacitación, ya que se requiere tiempo adicional comparado con PAPI.

¹⁰ Instituto de Investigaciones Económicas y Sociales, Universidad de Essex, Reino Unido

¹¹ Caeyers et al. (2010), prueba experimental realizada en Tanzania sobre un total de 1.840 hogares.

Los costos de impresión de cuestionario y digitación son prácticamente eliminados. Incluso eventualmente costos de clasificación y codificación. También, se producen ahorros significativos en la etapa de validación, dado que se acorta este proceso y el de edición.

Finalmente, la evidencia indica que CAPI conduce a mayores costos fijos iniciales, pero a menores costos variables por cuestionario, siendo a nivel agregado una forma de recolección de datos más económica y rápida que PAPI¹².

2. Medición de diferencia en estimaciones entre PAPI y CAPI

En el marco del proceso de levantamiento de Casen 2017, se realizaron dos estudios para evaluar la aplicación de la Encuesta Casen por medio de dispositivo móvil. Una de ellas fue realizada por el Centro de Microdatos (CMD) de la Universidad de Chile, y la segunda fue realizada por el Centro de Encuestas y Estudios Longitudinales (CEEL) de la Universidad Católica de Chile.

2.1. Metodología

En la Tabla 1 se presentan los principales antecedentes metodológicos de la realización de ambos estudios. Un elemento relevante para señalar es que los estudios difieren respecto al alcance territorial en que se realizaron las encuestas. El CMD realizó la encuesta en el área urbana y rural de la región Metropolitana, en tanto el CEEL la realizó en las regiones de Valparaíso, Metropolitana y Biobío en el área urbana y rural. Otra diferencia importante entre ambas pruebas es el tiempo destinado al terreno para cada modalidad y el tiempo total de ejecución del proyecto.

¹² Por ej. Leisher (2013), King et al. (2012) y Zhang et al. (2012) encontraron que los ahorros generados por el uso de CAPI fueron mayores que el desembolso inicial en hardware y software.

Tabla 1. Antecedentes del proceso de levantamiento entre las pruebas del CMD y del CEEL, año 2017

Prueba de campo Casen 2017	CMD		CEEL	
	Papel	Tablet	Papel	Tablet
N° muestra objetivo	950	950	1.100	1.100
N° entrevistas logradas	1.092	1.017	980	907
Periodo terreno	26 de abril al 27 de junio de 2017		3 de septiembre al 10 de octubre	18 de septiembre al 19 de octubre
N° de semanas terreno	9		5	4
N° encuestadores promedio semanal	23	21	31	32
Marco muestral	Hogares en viviendas encuestadas en Casen 2015		Censo de Población y Vivienda, año 2002	
Diseño muestral	Diseño cuasi experimental a nivel de hogares. El objetivo es identificar un cambio en la tasa de pobreza por ingresos superior a cuatro puntos porcentuales (se usa la tasa de pobreza de Casen 2015 en la Región Metropolitana, es decir, 7,1%). Se utiliza una submuestra estratificada de Casen 2015, de modo de contar con un doble grupo de control. La muestra es de la Región Metropolitana, zona urbana y rural e incluyó 2 muestras objetivos de 950 para PAPI y 950 para CAPI.		Diseño experimental por conglomerados a nivel de hogares. Este diseño incluyó 4 grupos experimentales de 550 hogares cada uno, para una muestra objetivo total de 2.200 viviendas (1.100 casos en PAPI y 1.100 casos en CAPI; 1.100 casos en Forma A y 1.100 casos en Forma B). La muestra es de las regiones de Valparaíso, Biobío y Metropolitana, zona urbana y rural.	
Metodología de Evaluación de Impacto	El método de análisis empleado es medir el efecto del tratamiento promedio (ATE) entre papel y Tablet. Además, se considera el método de diferencias en diferencias, utilizando la información de Casen 2015.		El método empleado es el análisis del efecto del tratamiento promedio (ATE) entre papel y Tablet.	

Fuente: Elaboración en base a Informe “Proyecto: Evaluación uso de Tablets en Encuesta Casen” del CMD e Informe “Reporte metodológico y de análisis de la Prueba de Campo de la Encuesta Casen 2017” del CEEL.

2.2. Resultados a nivel de procesos

2.2.1 Diseño e implementación

Tanto el CMD como el CEEL consideraron la contratación de personal informático para el desarrollo del cuestionario electrónico.

El CMD indica que el desarrollo de cuestionario electrónico tuvo una duración de 3 semanas y el formato utilizado fue matricial, es decir cada pregunta se responde simultáneamente por todos los integrantes del hogar a los que corresponda aplicar la pregunta. También se incorporó una lengüeta dinámica con las variables de nombre, sexo y edad con los objetivos de que: 1) el cuestionario electrónico sea similar al cuestionario en papel; 2) el encuestador cuente con esta información independiente de la pregunta en la cual se encuentre el encuestador. Además, una vez finalizada la programación del Tablet se contrató personal con experiencia en la encuesta Casen, para así poder realizar pruebas de validación de la programación.

Respecto de la capacitación de encuestadores, en ambas instituciones esta tuvo una duración de tres días, sin embargo, el CMD realizó una capacitación para los encuestadores Tablet y otra capacitación para los encuestadores Papel. Por otra parte, el CEEL realizó una capacitación para todos los encuestadores y una vez finalizada la capacitación se realizó la distribución de los encuestadores (PAPI/CAPI) mediante un diseño experimental factorial aleatorizado por conglomerados, con el objetivo de controlar la posible influencia de factores no observables asociados al encuestador.

2.2.2 Procesos operativos

El estudio realizado por el CMD concluyó que la duración promedio de la entrevista fue mayor en Tablet (66,3 min) que en Papel (54,9 min). Por otro lado, CEEL determinó que la duración promedio de la entrevista es mayor en Papel (59,0 min) que en Tablet (55,1 min).

Respecto del procesamiento de los datos, ambas pruebas evidencian que el hecho de realizar la encuesta en Tablet ayudó a que esta etapa fuera más rápida y menos costosa que en Papel, ya que la mayoría de las validaciones están consideradas en el aplicativo que contiene el cuestionario.

2.3. Medidas de calidad operativa

La tasa de no respuesta a la unidad y al ítem son considerados una medida de calidad operativa de la encuesta. Las instituciones responsables del levantamiento de ambas pruebas realizaron una serie de acciones con el objetivo de que estas medidas fueran óptimas como, por ejemplo: reclutamiento de encuestadores con experiencia en levantamientos previos de Casen, visitar direcciones en al menos tres oportunidades, uso de accesorios institucionales asociados a la respectiva universidad con el fin de aumentar la cooperación y disminuir rechazos, entre otros.

2.3.1 No respuesta a la Unidad

En la Tabla 2 se presentan los indicadores de resultados de trabajo de campo, tanto para el CMD como el CEEL, para ambas modalidades de aplicación. Se observa que en la prueba realizada por el CMD los indicadores mostraron mejores tasas en la modalidad Papel, en tanto en la aplicación realizada por el CEEL, los indicadores mostraron mejores tasas en la modalidad Tablet.

Tabla 2. Resultado del Trabajo Campo por modalidad de aplicación según indicador, pruebas CMD y CEEL, año 2017 (%)

Indicador	CMD		CEEL	
	Tablet	Papel	Tablet	Papel
Tasa respuesta	62,2	66,8	60,0	54,9
Tasa cooperación	89,7	93,5	94,3	89,0
Tasa rechazo	5,2	3,6	3,4	6,3
Tasa contacto	69,4	71,4	63,6	61,0

Fuente: Elaboración en base a Informe “Proyecto: Evaluación uso de Tablets en Encuesta Casen” del CMD e Informe “Reporte metodológico y de análisis de la Prueba de Campo de la Encuesta Casen 2017” del CEEL.

2.3.2 No respuesta al Ítem

Se entiende por no respuesta al ítem cuando el informante, ante cualquiera de las preguntas de la encuesta, señala no saber la respuesta o prefiere no responder.

En la Tabla 3 se presentan las tasas de no respuesta al ítem de las corrientes de ingresos que pasan por el proceso de corrección definido por la Cepal¹³. Para el caso de CEEL, las tasas de no respuesta al ítem para las variables de ingresos analizadas en la modalidad Tablet son significativamente mayores respecto a la modalidad Papel, con la excepción de la pregunta de ingresos por jubilación. Por otro lado, en la prueba realizada por el CMD, sólo se presentan diferencias estadísticamente significativas en las variables y1 y v22, siendo mayores las tasas de no respuesta en Tablet respecto a Papel.

¹³ Más detalle del proceso de corrección se encuentra en el documento “Medición de los ingresos y la pobreza en Chile, Encuesta Casen 2017”, disponible en http://observatorio.ministeriodesarrollosocial.gob.cl/storage/docs/Casen/2017/Medicion de la pobreza en Chile 2017_17082018.pdf

Tabla 3. Tasa de no respuesta al ítem por modalidad de aplicación según tipo de ingreso, pruebas CMD y CEEL, año 2017 (%)

Tipo de ingreso	CMD		CEEL	
	Tablet	Papel	Tablet	Papel
y1. Ingreso ocupación principal-Asalariados	6,5**	4,5	12,4**	5,8
y7. Ingreso ocupación principal- No asalariados	6,9	7,2	13,2**	5,8
y26_2c. Ingresos por jubilación	5,1	3,5	6,9	3,5
v22 ¹⁴ . Alquiler imputado	6,7**	1,3	31,4**	1,3

Fuente: Elaboración propia según bases de datos Piloto realizadas por el CMD y CEEL.

Nota**: Significancia estadística al 95% de confianza. Los valores corresponden a pruebas de hipótesis de proporciones para dos muestras independientes, calculado como la diferencia entre P(Tablet) – P(Papel).

2.4. Resultados a nivel de estimaciones de los indicadores principales de la encuesta¹⁵

A continuación, se presentan los principales indicadores correspondientes a la tasa de pobreza por ingreso, pobreza multidimensional y los promedios de ingresos de las principales corrientes de ingreso.

2.4.1 Tasas de pobreza

En la Tabla 4 se muestra la situación de pobreza por ingresos y multidimensional a nivel de hogar y personas. Los datos del CMD muestran que las tasas de pobreza por ingresos y multidimensional no presentan diferencias estadísticamente significativas por modalidad de aplicación. En tanto, los resultados de la prueba desarrollada por CEEL muestran que la tasa de pobreza multidimensional, a nivel de personas, es significativamente mayor en la modalidad Tablet. En los otros resultados de la prueba de CEEL, las diferencias observadas no son estadísticamente significativas al 95% de confianza.

Tabla 4. Indicadores de pobreza por modalidad de aplicación, pruebas CMD y CEEL, año 2017 (%)

Indicador	CMD		CEEL	
	Tablet	Papel	Tablet	Papel
Tasa pobreza ingresos-Hogar	4,8	4,8	8,3	8,9
Tasa pobreza ingresos-Personas	5,1	5,2	9,2	10,4
Tasa pobreza -multidimensional-Hogar	14,5	14,1	17,8	16,5
Tasa pobreza -multidimensional-Personas	19,1	18,3	23,5**	20,9

Fuente: Elaboración propia según bases de datos Piloto realizadas por el CMD y CEEL.

Nota**: Significancia estadística al 95% de confianza. Los test corresponden a pruebas de hipótesis de proporciones para dos muestras independientes, calculado como la diferencia entre P(Tablet) – P(Papel).

¹⁴ Por definición el alquiler imputado aplica solamente a quienes son propietarios de su vivienda (incluyendo cesiones y usufructo). Por este motivo, no se consideró en el análisis las categorías “arrendada con contrato”, “arrendada sin contrato”, “ocupación irregular (de hecho)” y “poseedor irregular”.

¹⁵ Los resultados presentados no son comparables entre CMD y CEEL, dadas las características del diseño muestral. Por tanto, las comparaciones se realizan según modo de aplicación en cada institución.

2.4.2 Ingresos

En la Tabla 5 se detallan los ingresos promedios¹⁶ según modalidad de aplicación. Los resultados de la prueba desarrollada por el CMD muestran que los ingresos en análisis no presentan diferencias significativas por modalidad. Por otro lado, los resultados de la prueba piloto desarrollada por el CEEL muestran que el ingreso autónomo del hogar y los subsidios del hogar presentan diferencias significativas al 95% de confianza por modalidad, mientras que en el ingreso total del hogar y en el alquiler imputado no se observan diferencias significativas.

Tabla 5: Ingreso promedio del hogar por modalidad de aplicación según tipo de ingreso, pruebas CMD y CEEL, año 2017

Tipo de ingreso	Modo	CMD	CEEL
		Promedio (\$)	Promedio (\$)
Ingreso autónomo del hogar-correctado	Papel	1.039.352	746.727**
	Tablet	1.128.134	683.877
Subsidios del hogar	Papel	30.199	31.513
	Tablet	32.474	37.680**
Ingreso total del hogar-correctado	Papel	1.294.539	948.617
	Tablet	1.379.822	891.143
Alquiler imputado del hogar-correctado	Papel	224.988	170.378
	Tablet	219.214	169.585

Fuente: Elaboración propia según bases de datos Piloto realizadas por el CMD y CEEL.

Nota**: Significancia estadística al 95% de confianza. Los valores corresponden a pruebas de hipótesis de medias para dos muestras independientes, calculado como la diferencia entre P(Papel) – P(Tablet).

2.5. Conclusiones y recomendaciones realizadas por CMD y CEEL

Ambas instituciones recomiendan la aplicación de la encuesta Casen a través de dispositivo móvil (CAPI) en futuros levantamientos. Las principales conclusiones de las pruebas informadas por ambas instituciones van en la línea de las recomendaciones y experiencias internacionales, y se pueden resumir en:

- 1) El procesamiento de datos es más rápido y menos costoso que Papel.
- 2) No se requiere equipo exclusivo de validación, digitación o lectura óptica.
- 3) Menores o nulos costos de impresión de cuestionarios, manuales, hoja de rutas entre otros.
- 4) Menores o nulos costos asociados al transporte de material.
- 5) Menos o nulos costos asociados a la seguridad o resguardo de material.

Además, señalan que la decisión a favor de la aplicación en dispositivo móvil debe adoptarse no mirando los resultados o estimaciones en indicadores de pobreza o desigualdad de ingresos entre las

¹⁶ Los ingresos consideran corrección de ingresos siguiendo la metodología desarrollada por Cepal. Ver: http://observatorio.ministeriodesarrollosocial.gob.cl/storage/docs/Casen/2017/Medicion_de_la_pobreza_en_Chile_2017_17082018.pdf

dos modalidades, sino sobre la base de la evidencia en la mejora de los procesos, la calidad de los datos y de la oportunidad en la entrega de la información.

3. CAPI en Casen en Pandemia 2020

Para efectos del desarrollo de la versión CAPI del cuestionario de Casen en Pandemia 2020, se usó como base la versión del piloto del año 2017 que fue revisada para el proceso 2019. Todas estas fueron desarrolladas por el CEEL (proveedor de Casen en Pandemia 2020).

En primera instancia, se procedió a incorporar todos los cambios de cuestionario de la versión 2019¹⁷. En los inicios del proyecto (fase “alfa”), se considera la adaptación de preguntas a CAPI, la incorporación de filtros y universos, y la incorporación de validaciones que se hacían habitualmente en las etapas de procesamiento, entre otros.

Posteriormente, se avanzó hacia una fase beta que dio origen a la versión 1.0 del aplicativo. A partir de las pruebas internas del CEEL y del Ministerio, se realizaron una serie de observaciones al aplicativo. Esto inició un ciclo de desarrollo, donde aproximadamente cada dos o tres semanas, se lanzaba una nueva versión habilitada para pruebas. Este período se extendió por aproximadamente 2 meses.

A continuación, se detallan las principales características del aplicativo CAPI, los desarrollos de programación y una breve descripción del historial de cambios en cada versión.

3.1. Lógica de programación del aplicativo

Una de las características claves del desarrollo del aplicativo del CEEL guarda relación con su lógica de programación. En particular, el software busca las respuestas correctas, o lo que se define de antemano como “verdadero”. Es decir, dado un input de programación, el software verifica si este es correcto, permitiendo el paso. En principio, cuando no se cumple esta condición, es decir, cuando el software evalúa la condición lógica como “falsa”, se entrega un mensaje estándar indicando el error.

Históricamente, para los procesos de validación se ha usado como base una malla de validación que contiene el listado de todos los errores posibles. Cuando se recibe la base de datos se aplicaba esta malla y se depuraban las inconsistencias. La lógica de programación antes descrita trata de emular lo que tradicionalmente se ha desarrollado en versiones previas de la Encuesta Casen en papel.

3.2. Tipos de preguntas y formatos

El formato CAPI posibilitó una serie de mejoras al cuestionario Casen, en comparación al cuestionario en papel. Una de ellas fue la posibilidad de estandarizar el formato de aplicación de las preguntas¹⁸. Se optó por dos alternativas, disposición lineal y disposición matricial:

¹⁷ La Casen 2019 fue suspendida debido a la crisis social acaecida a partir de octubre de 2019, pero el cuestionario ya había sido definido.

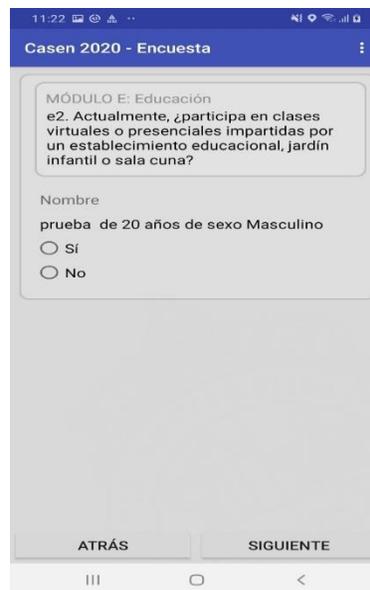
¹⁸ En formato papel, el encuestador puede aplicar las preguntas en formato lineal (a cada persona un determinado flujo de preguntas) o matricial (la misma pregunta a todas las personas del hogar a quienes aplica dicha pregunta).

- Formato lineal: corresponde a preguntas del cuestionario que se aplican de manera sucesiva a un mismo sujeto. Una vez finalizado ese grupo de preguntas, se continúa con la siguiente persona, en caso de que aplique. Por ejemplo, en el módulo de Salud, se comienza con el primer encuestado y se le aplica el módulo completo de preguntas. Posteriormente, se continúa con la siguiente persona del listado.
- Formato matricial: corresponden a grupos de preguntas que se aplican de una vez a todos los integrantes del hogar seleccionables. Así, a modo de ejemplo, desde las preguntas y12 a y25 del módulo de Ingresos se consulta simultáneamente por todas las personas a las que corresponde aplicar la pregunta.

Paralelamente a la definición del formato matricial o lineal, se debió definir el “tipo de objeto”. A grandes rasgos, se puede diferenciar entre los siguientes.

1. **Preguntas de respuesta única con botón seleccionable:** corresponden a preguntas aplicadas de manera lineal y de respuesta única.

Figura 1. Visualización de tipo de pregunta e2, Modulo Educación



The screenshot shows a mobile application interface for the 'Casen 2020 - Encuesta'. At the top, the status bar shows the time 11:22 and various icons. Below the status bar, the title 'Casen 2020 - Encuesta' is displayed in a blue header. The main content area is titled 'MÓDULO E: Educación' and contains the question: 'e2. Actualmente, ¿participa en clases virtuales o presenciales impartidas por un establecimiento educacional, jardín infantil o sala cuna?'. Below the question, the user's name is displayed as 'Nombre prueba de 20 años de sexo Masculino'. There are two radio button options: 'Sí' and 'No'. At the bottom of the screen, there are two buttons: 'ATRÁS' and 'SIGUIENTE'. The Android navigation bar is visible at the very bottom.

Fuente: Captura de pantalla aplicativo de recolección de datos Casen en Pandemia 2020.

2. **Preguntas de respuesta abierta:** permiten la inclusión de texto alfanumérico, con la posibilidad de limitar el número de caracteres. Adicionalmente, también es posible limitar entre teclado numérico y alfanumérico.

Figura 2. Visualización de tipo de pregunta o9b, Modulo Trabajo

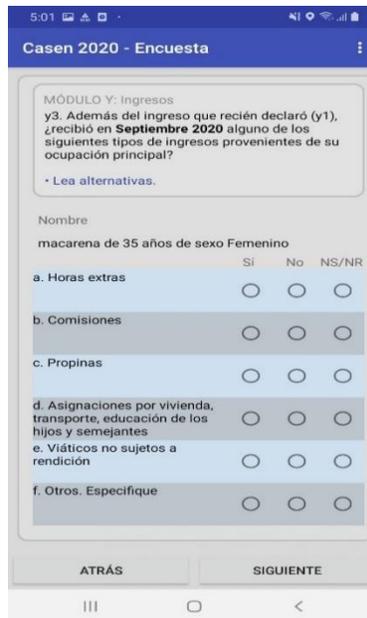


The screenshot shows a mobile application interface for the 'Casen 2020 - Encuesta'. The question is: 'o9b. ¿Que hace en su trabajo o negocio principal?'. Below the question, there is a list of examples: 'Recoger y transportar uvas', 'Enseñar matemáticas', 'Cuidar enfermos y administrar medicamentos', 'Transportar pasajeros entre ciudades', 'Vigilar y controlar entrada en una empresa', and 'Administrar una empresa de productos lácteos'. The user's name is 'manual de 20 años de sexo Masculino'. There is a text input field labeled 'Describa' and two buttons: 'ATRÁS' and 'SIGUIENTE'. A keyboard is visible at the bottom of the screen.

Fuente: Captura de pantalla aplicativo de recolección de datos Casen en Pandemia 2020.

3. **Preguntas “radio button” en formato lineal:** corresponde a preguntas con un círculo de activación donde cada fila corresponde a un ítem específico. Cabe mencionar que, como cada ítem comparte el mismo enunciado de pregunta, este tipo de objeto optimiza la inclusión de más preguntas en una misma pantalla.

Figura 3. Visualización de tipo de pregunta y3, Modulo Ingresos

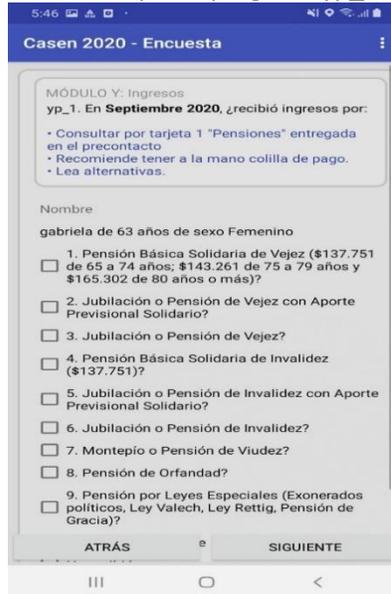


The screenshot shows a mobile application interface for the 'Casen 2020 - Encuesta'. The question is: 'MÓDULO Y: Ingresos y3. Además del ingreso que recién declaró (y1), ¿recibió en Septiembre 2020 alguno de los siguientes tipos de ingresos provenientes de su ocupación principal?'. Below the question, there is a list of income types: 'Horas extras', 'Comisiones', 'Propinas', 'Asignaciones por vivienda, transporte, educación de los hijos y semejantes', 'Viáticos no sujetos a rendición', and 'Otros. Especifique'. Each item has three radio buttons labeled 'SI', 'No', and 'NS/NR'. The user's name is 'macarena de 35 años de sexo Femenino'. There are two buttons: 'ATRÁS' and 'SIGUIENTE'.

Fuente: Captura de pantalla aplicativo de recolección de datos Casen en Pandemia 2020.

4. **Preguntas de selección múltiple:** este tipo de objeto permite la selección de más de una alternativa de respuestas, con la posibilidad de especificar la cantidad de respuestas. Además, como se aprecia en el ejemplo, algunas de estas alternativas solicitan especificar con más información a partir de texto abierto.

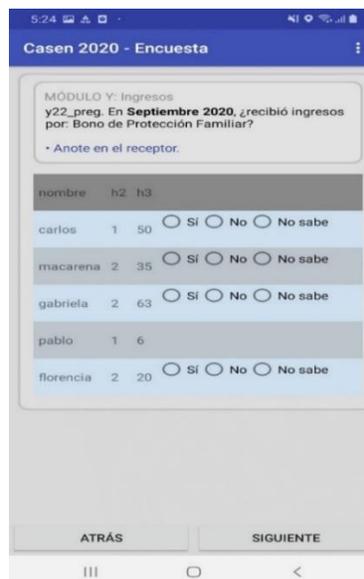
Figura 4. Visualización de tipo de pregunta yp_1, Modulo Ingresos



Fuente: Captura de pantalla aplicativo de recolección de datos Casen en Pandemia 2020.

5. **Preguntas “radio button” matricial:** en este caso, cada circulo de activación corresponde a las alternativas de respuesta y las filas a los encuestados.

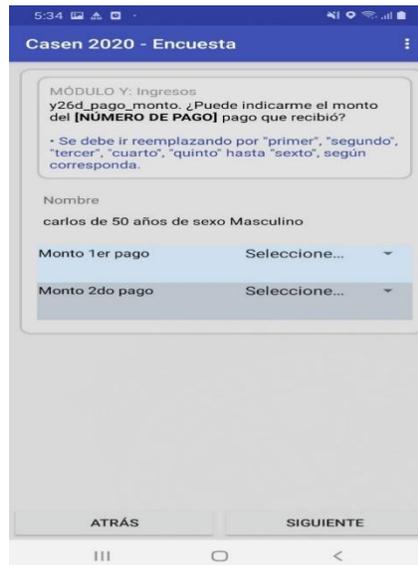
Figura 5. Visualización de tipo de pregunta y22_preg, Modulo Ingresos



Fuente: Captura de pantalla aplicativo de recolección de datos Casen en Pandemia 2020.

6. **Preguntas con menú desplegable:** este tipo de objeto es utilizado para las preguntas matriciales. Las alternativas de respuesta están contenidas en un menú desplegable. En contraste al *radio button* matricial, este tipo se utiliza para alternativas de respuesta que tienen mayor longitud de texto.

Figura 6. Visualización de tipo de pregunta y26d, Modulo Ingresos



Fuente: Captura de pantalla aplicativo de recolección de datos Casen en Pandemia 2020.

3.3. Opciones de programación en CAPI Casen en Pandemia 2020

En adición a los tipos de disposición de preguntas y objetos, la programación CAPI desarrollada por CEEL ha incorporado el siguiente listado de funcionalidades que se muestra en la Tabla 6.

Tabla 6. Descripción de las funcionalidades implementadas en Casen en Pandemia 2020

Funcionalidad	Descripción	Tipo de función
Validación automática mediante flujo del cuestionario	La programación del aplicativo evita una serie de errores que se producen en PAPI producto de aplicaciones erróneas de flujo de cuestionario.	Validación de cuestionario
Validación entre respuestas mediante mensaje de error	Cuando se produce una inconsistencia entre alternativas de respuesta, el dispositivo entrega un mensaje que alerta del error.	Validación de cuestionario
Ocultar alternativas de respuesta según flujo del cuestionario	El aplicativo permite ocultar determinadas alternativas de respuesta que no apliquen para grupos específicos o según lo respondido previamente	Validación de cuestionario
Orientación vertical u horizontal	El aplicativo permite la disposición de manera vertical u horizontal, según preferencia del encuestador.	Usabilidad
Texto dinámico	Para las preguntas que refieren al nombre de la persona, el software indicará el nombre específico y la edad de quien se está consultando. Por otro lado, para las preguntas con periodo temporal del módulo de ingresos, el campo “el mes pasado” se modifica automáticamente según el periodo en que fue realizada la encuesta.	Usabilidad

Funcionalidad	Descripción	Tipo de función
Autocompletado de texto	Para las preguntas que refieren a países o comunas, una vez el encuestador va escribiendo, el programa muestra el listado para seleccionar.	Usabilidad
Texto multilínea	El aplicativo permite textos en más de una línea.	Usabilidad
Retroceder sin perder respuesta	El aplicativo permite retroceder sin perder respuestas.	Usabilidad
Ayuda	El programa tiene la opción de definir mensajes de ayuda para cada pregunta mediante un ícono que, al presionar, despliega información en pantalla.	Usabilidad
Retomar encuestas	El aplicativo permite retomar en cualquier momento una encuesta pendiente.	Usabilidad
Geolocalización	Mediante Google Maps, es posible saber el punto geográfico donde se realizó una encuesta específica.	Monitoreo de trabajo de campo
Cámara fotográfica	Se incluye en la finalización de la encuesta, la factibilidad de sacar una fotografía o fachada de la casa o edificio, a modo de prueba del trabajo realizado	Monitoreo de trabajo de campo
Errores suaves	Para patrones de respuesta poco usuales o valores fuera de rango, el aplicativo solicita confirmar la respuesta por parte del encuestador	Calidad de los datos
Núcleos	Gracias a la programación CAPI, se diseñó una nueva metodología de registro de los núcleos familiares. Esto implicó ajustes en las preguntas para que no quedaran definiciones a discrecionalidad del encuestador sino a cargo del aplicativo, diseñando nuevos formatos de pregunta y programando la creación de los núcleos.	Calidad de los datos
Mensajes para no respuesta	En preguntas clave, cuando se marca la alternativa "no sabe", el programa permite entrega un mensaje que solicita confirmar respuesta	Calidad de los datos
Resumen error en pantalla	Específicamente para el módulo de Registro se habilita un cuadro resumen de errores suaves	Calidad de los datos

Fuente: Elaboración propia

4. Usabilidad y Aseguramiento de Calidad del aplicativo Casen en Pandemia 2020

Según la norma ISO/IEC 9241, la “usabilidad” es la eficacia, eficiencia y satisfacción con la que un producto permite alcanzar objetivos específicos a usuarios específicos en un contexto de uso específico, que en el caso de un aplicativo móvil corresponde a la interfaz digital mediante (o con ayuda de) la cual se busca realizar una tarea.

Por otra parte, Nielsen (2012) la define como un atributo de calidad que evalúa la facilidad de uso de las interfaces de usuario. La palabra "usabilidad" también se refiere a métodos para mejorar la facilidad de uso durante el proceso de diseño.

Estas definiciones se complementan, permitiendo entender que usabilidad es tanto un atributo que involucra la experiencia individual como el logro de una tarea, a la vez que hace referencia a un proceso de diseño dentro del desarrollo de plataformas digitales.

Las pruebas de usabilidad se realizan entre otros motivos para identificar en qué medida una interfaz es fácil de usar. A mayor grado de usabilidad el usuario puede realizar de manera eficaz, eficiente e intuitiva la tarea que le ha sido asignada.

Otra característica relevante de los test de usabilidad es que estos son mayoritariamente cualitativos. Se basan principalmente en la observación del usuario al interactuar con la interfaz y en las respuestas que da a preguntas basadas en la experiencia con ella.

En el caso de las encuestas, el paso de una modalidad en papel a una digital significa un cambio significativo para los distintos usuarios involucrados en la gestión y ejecución del proyecto.

4.1. Prueba de usabilidad realizadas por el CEEL

Las pruebas de usabilidad tienen puesto el foco en los usuarios. En este sentido, en Casen en Pandemia 2020 se definen dos tipos de usuario: los usuarios del software de recolección de datos, es decir encuestadores y supervisores (entendiendo por estos últimos a los encuestadores de las encuestas de supervisión interna); y, por otra, de los usuarios del software de los sistemas web de seguimiento de la muestra y control de calidad, a saber, los coordinadores de grupo y de sede.

De lo anterior se desprende que las plataformas (y usuarios) consideradas relevantes de testear son el Software de Sistema de Control y Seguimiento de la Muestra/ Software de Sistema de Control de Calidad en Terreno (web) cuyos usuarios son: Coordinadores de Grupo y Coordinadores de Zona; y el Software de Recolección de Datos (aplicativo móvil) cuyos usuarios son Encuestadores y Supervisores.

Las pruebas de usabilidad son consideradas como un método fundamentalmente cualitativo para evaluar una interfaz digital, por lo tanto, la muestra que realizará el test no debe ser estadísticamente representativa. Nielsen (2012), señala que testear una plataforma digital con 5 usuarios es más que suficiente, ya que permite hacer mejoras de manera ágil y eficiente. Para este caso se considera relevante que mediante el test de usabilidad se abarque el mayor espectro posible de características

personales que puedan ser eventualmente conflictivas al momento de utilizar las interfaces digitales en la aplicación de la encuesta Casen.

En este sentido, se diseñaron pruebas considerando perfiles de usuarios de acuerdo a la edad (“joven” y “mayor”), y experiencia en Casen (“primerizo” o “experimentado”). Para cada uno de los perfiles se presentaron diferentes escenarios que contextualizan al usuario dentro de casos hipotéticos. Posteriormente, en cada una de las dimensiones evaluadas, se construyeron indicadores asociados al nivel de logro de la tarea, el tiempo demorado en completarla, el número de errores, y la facilidad de uso.

Con estas actividades, además de preguntas realizadas a los usuarios, se obtuvieron resultados para entender la experiencia con el dispositivo móvil, así como la migración de la encuesta desde el papel a lo digital.

En relación con el traspaso del papel a la aplicación digital, los usuarios que participaron en la prueba de usabilidad tenían en su mayoría experiencia previa con encuestas PAPI (y algunos además en Casen), por lo que resultó relevante el hallazgo que la aplicación fue un elemento positivo y facilitador del proceso de encuesta, pero que necesitaban aprender a utilizarla bien, lo que se relaciona con el proceso de capacitación de los encuestadores Casen. En resumen, una buena capacitación, preguntas claramente redactadas y una interfaz usable son los elementos clave para que los encuestadores puedan hacer un levantamiento correcto de datos en terreno.

De estas pruebas de usabilidad se levantaron hallazgos, con diferentes niveles de criticidad, que fueron abordados en el desarrollo de las versiones sucesivas del aplicativo de recolección de datos, algunas de ellas a nivel de preguntas (conceptos necesarios a reforzar en las capacitaciones) y otros netamente de usabilidad (tamaño de letras, extensión de textos, resaltar ciertos textos, códigos de las categorías).

Tabla 7. Resultados pruebas de usabilidad CEEL - Casen en Pandemia 2020

<u>Niveles de Criticidad</u>
Criticidad Alta: <ol style="list-style-type: none">1. Códigos de disposición final de casos generan confusión para la mayoría de los encuestadores.2. Concepto de núcleo es complejo.3. Existen Instrucciones muy largas y/o poco claras.
Criticidad Media: <ol style="list-style-type: none">1. Usuarios olvidan sincronizar la app.2. Ícono para entrar a encuesta no es autoexplicativo.3. Botones de Agregar Persona y Continuar son iguales y pueden inducir error.4. Confusión porque solo se pide información de menores de edad.5. Diferencias entre el menú desplegable y tarjetero genera riesgo de error6. Usuarios olvidan o no usan tarjetero.7. Texto de instrucciones no es coherente en su forma.8. Presencia de lenguaje de programación o técnico que puede causar confusión.
Criticidad Baja: <ol style="list-style-type: none">1. Tamaño de la letra (no, dado que no se considera relevante).2. Pantalla Horizontal genera incomodidad (no, dado que no se considera relevante).3. Tamaño del área táctil en elementos interactivos puede ser una dificultad para algunos usuarios (no, dado que no se considera relevante).

Fuente: Elaborado por CEEL UC, Casen en Pandemia 2020.

Posteriormente, al alcanzar una versión más estable del aplicativo de recolección de datos, el contratista realizó una evaluación heurística de la usabilidad del aplicativo. Esta actividad se realizó mediante una revisión de cada una de las pantallas existentes, además de la navegación para la simulación de casos ficticios de la encuesta.

El criterio en base al cual se realizó la evaluación del software de recolección de datos consideró los principios de usabilidad establecidos por dos autores:

- Por una parte, se consideraron los 10 principios de heurística planteados por Jakob Nielsen (ver Nielsen, 2019):
 1. Visibilidad del estado del sistema
 2. Utilizar el mismo lenguaje que el usuario
 3. Control y libertad para el usuario
 4. Consistencia y estándares
 5. Prevención de errores
 6. Minimizar la carga de memoria del usuario
 7. Flexibilidad y eficiencia de uso
 8. Diálogos estéticos y diseño minimalista

9. Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de los errores

10. Ayuda y documentación

- Por otra parte, se consideraron los 8 principios heurísticos propuestos por Ben Shneiderman (ver Shneiderman et al, 2019):
 1. Trabaja por la consistencia
 2. Permite que los usuarios frecuentes utilicen atajos
 3. Ofrece comentarios informativos
 4. Diseña un diálogo para notificar la conclusión
 5. Ofrece un manejo simple de errores
 6. Permite una reversión fácil de acciones
 7. Soporta un punto de control interno
 8. Reduce la carga de memoria a corto plazo

En base a esto, se hizo una revisión del dispositivo móvil de acuerdo con el contexto de uso y sus usuarios. Para esto se observó una a una las pantallas de manera estática y en uso, a partir de lo cual se extrajeron observaciones. Posteriormente, las observaciones fueron agrupadas en base a la temática que abordaban, a partir de lo cual se generaron 2 grupos:

- Observaciones de pantallas particulares: Observaciones a pantallas que por su especificidad y/o formato no compartían observaciones con otras.
- Observaciones Transversales: Observaciones que se aplican a múltiples pantallas.

Además, se definió que para cada tipo de observación se asignaría una recomendación, con el fin de orientar el proceso de mejora continua en el desarrollo de esta y futuras encuestas.

De este proceso de revisión, se concluyó que ninguna de las observaciones hechas compromete la eficacia en el levantamiento de información. De esta manera, las observaciones levantadas condujeron a mejoras implementadas finalmente en la versión utilizada en el trabajo de campo.

4.2. Pruebas de Aseguramiento de la Calidad (QA) del Software

La prueba de aseguramiento de calidad consiste en la revisión y certificación del correcto funcionamiento de aplicaciones y sistemas, ya sea en un ambiente cerrado o interactuando con el resto del entorno en donde finalmente operará, asegurando el cumplimiento del nivel de calidad esperado por el usuario. Esto se logra a través de la ejecución de pruebas funcionales, de ciclo cero, de no impacto, de código, de seguridad, de rendimiento y de estrés.

Para estos fines, se aplicaron pruebas de QA por parte de CEEL y por parte del Ministerio, cada una con objetivos específicos que permitían complementar la revisión y asegurar el correcto funcionamiento del software.

A continuación, se describe brevemente los aspectos que evalúan cada una de estas pruebas:

Pruebas de QA programadas:

a) Pruebas de vulnerabilidad

Una vulnerabilidad es una debilidad o fallo en un sistema de información que pone en riesgo la seguridad de la información pudiendo permitir que un atacante pueda comprometer la integridad, disponibilidad o confidencialidad de la misma, por lo que es necesario encontrarlas y eliminarlas lo antes posible. Estos «agujeros» pueden tener distintos orígenes, por ejemplo: fallos de diseño, errores de configuración o carencias de procedimientos.

b) Pruebas Unitarias

Para corroborar el correcto funcionamiento de cada una de las unidades de código y verificar sus resultados, se realizaron pruebas unitarias o test unitarios. Además de verificar que el código hace lo que tiene que hacer en cada una de sus unidades, se verifica que sean correctos los nombres y tipos parámetros, el tipo de resultado que se devuelve, que, si el estado inicial es válido, entonces el estado final también debe ser válido.

c) Prueba de carga y estrés

Para asegurar la disponibilidad de los distintos sistemas, se realizaron pruebas de carga para validar la respuesta de la aplicación cuando fuera sometida a una gran cantidad de usuarios y/o transacciones simultáneas que se espera en el ambiente de producción. Esta prueba fue realizada tanto para la plataforma de recepción de datos del dispositivo móvil (Servicio web) como para la página web de gestión. Las pruebas de estrés sirven para determinar cuál es la cantidad de datos que se necesita para que el sistema o aplicación comience a fallar o sea incapaz de responder a estas peticiones superando los límites esperados en el ambiente de producción.

d) Pruebas de Integración

El objetivo fue recoger la definición de las pruebas que deberán realizarse para asegurar la correcta integración entre todos los componentes/módulos del sistema.

e) Pruebas funcionales

Para este tipo de pruebas se utilizó Visual Studio 2017, el cual permite realizar pruebas Codificadas UI (CUIT) o pruebas de interfaz de usuario. Xamarin. UITest es un framework de testing que permite realizar pruebas de comportamiento de la aplicación automatizando interacciones con la misma. Integrado con Xamarin.Android, permite automatizar la gestión con la interfaz con pulsación de botones, introducir textos, gestos, tomar capturas, etc. Cada UITest es un método escrito en C# que sigue el patrón Arrange-Act-Assert: el test inicializa todo lo necesario para que pueda ser lanzado. También es una buena forma de automatizar tareas repetitivas cuando se trabaja con el código de la interfaz de usuario.

Estas pruebas fueron aplicadas a todos los componentes, tales como el Aplicativo de recolección de datos, Sistema de control y seguimiento de la muestra y Sistema de control de calidad en terreno.

f) Pruebas de aceptación

El objetivo fue definir el conjunto de pruebas que deberán ser ejecutadas por los usuarios del sistema para validar si el sistema cumple con los requisitos de funcionamiento esperado y proceder así a la aceptación del sistema. Estas pruebas fueron dirigidas a los encuestadores, coordinadores y supervisores usuarios de los distintos sistemas a utilizar (Aplicativo de recolección de datos, Sistema de control y seguimiento de la muestra, Sistema de control de calidad en terreno).

g) Análisis de código

Para el análisis de código se utilizó Visual Studio 2017 que viene con un analizador de código incorporado. Los analizadores de .NET Compiler Platform ("Roslyn") analizan el estilo, la calidad, el mantenimiento y el diseño del código, además de otros problemas. Visual Studio incluye un conjunto integrado de analizadores que analizan el código de C# o Visual Basic a medida que escribe. Si un analizador detectaba infracciones de reglas, se notificaban en el editor de código (como un subrayado ondulado bajo el código infractor) y en la ventana Lista de errores.

h) Seguridad de Red

Para el aseguramiento de los datos, el Data Center utilizado contaba con un Firewall el cual protege los recursos de los servidores, el cual solo dispuso el puerto 80 de acceso al servidor de aplicaciones.

Los equipos internos se encontraban en una red cerrada sin acceso desde el exterior, siendo el único acceso a través de conexión VPN por el software Cisco System VPN Client. Esta VPN es una red privada virtual (RPV), en inglés: Virtual Private Network (VPN), es una tecnología de red de computadoras que permite una extensión segura de la red de área local (LAN) sobre una red pública o no controlada como Internet. Permite que la computadora en la red envíe y reciba datos sobre redes compartidas o públicas como si fuera una red privada con toda la funcionalidad, seguridad y políticas de gestión de una red privada.

i) Actualizaciones Críticas

El sistema operativo destinado para los servidores con los diferentes servicios para el proyecto Casen en Pandemia 2020 Microsoft Windows Server 2012 R2, para las actualizaciones relacionadas con la seguridad o no, es a través de Windows Update. Estas actualizaciones solucionan problemas y mejoran la confiabilidad en general del sistema operativo. Estas actualizaciones serán programadas de acuerdo a la necesidad y el riesgo que tengan. También se considera el mismo criterio para los dispositivos móviles con el sistema operativo Android. Para este tipo de actualizaciones se definió un protocolo de actualización que no pusiera en riesgo la continuidad operativa del proyecto.

Estas 9 pruebas fueron ejecutadas en diferentes etapas del desarrollo de los componentes del software (aplicativo de recolección de datos, sistema de control y seguimiento de la muestra y sistema de control de calidad).

Los resultados de estas pruebas fueron reportados en diferentes informes por parte del contratista, sin dar cuenta de errores críticos que pusieran en riesgo el desarrollo del proyecto desde un punto de vista técnico. En la revisión de las incidencias reportadas, así como de las correcciones desarrolladas por parte de contratista, participó el equipo de la División del Información Social del Ministerio, quienes desarrollaron funciones de contraparte técnica en la revisión de estas pruebas.

Adicionalmente, el plan de pruebas de aseguramiento de calidad del Ministerio de Desarrollo Social y Familia se ejecutó como una tarea conjunta entre la División Observatorio Social y la División de Información Social, ambas de la Subsecretaría de Evaluación Social. Estas dos áreas funcionaban en la práctica de forma integrada, puesto que cualquier regla incorporada o modificada en la malla de validación de errores, debía ser programada en el aplicativo de recolección de datos y luego se debía revisar su correcta implementación en las pruebas manuales o funcionales.

Malla de Validación

El proceso de elaboración de la malla de validación de errores sigue el proceso que se ha utilizado en versiones anteriores de la encuesta Casen, con algunas mejoras de acuerdo al cambio en la metodología de aplicación (CAPI).

Tradicionalmente en la malla de validación de errores, que se aplicaba una vez los datos eran traspasados del papel a un medio digital (base de datos), se reportaban errores de rango y consistencia. La programación CAPI incorpora gran parte de los errores dentro del flujo del cuestionario, por lo que los errores se definieron en función del tipo de alerta o restricción que se quiere generar. En base a esto, se definieron 3 tipos de errores: estrictos, suaves, y ex post. Los dos primeros incorporados en la misma aplicación, mientras que el tercero se revisaba una vez que la información se cargaba en la base de datos.

- **Errores estrictos:** Estos errores establecen condiciones que se deben cumplir en el flujo del cuestionario para poder avanzar a la siguiente pregunta. En este tipo de errores entran por ejemplo los de universo, donde si una determinada pregunta era aplicada a un universo específica, y una persona con esas características no registraba una respuesta en dicha pregunta no se podía avanzar a la siguiente.
- **Errores suaves:** Este tipo de errores estaban diseñados para generar una alerta en la aplicación, para que posterior a esta, el encuestador pudiera confirmar el dato registrado y seguir en el flujo del cuestionario. Este tipo de errores se implementó en mayor medida en las preguntas de ingresos, en función de los valores fuera de los rangos validos implementados en la aplicación.
- **Errores Ex post:** Este tipo de error fue diseñado para revisar casos en que, por la complejidad de la información, deben ser evaluados con mayor detalle por el equipo de validación. Un ejemplo de este tipo de error es cuando una persona declara en el módulo de trabajo que la

razón por no trabajar ni buscar trabajo en la última semana es por encontrarse estudiando (o7=11), pero en el módulo de educación declaró que no asiste a un Establecimiento Educativo (e2=2).

Pruebas manuales o QA funcional

En cuanto al proceso de QA funcional, su objetivo principal fue determinar la estabilidad de la aplicación móvil, comenzado por la realización de encuestas y considerando la realización de pruebas unitarias, integradas y flujos de ejecución de principio a fin. Para esto, se crearon casos tipo que buscaban intencionar errores en la aplicación.

El proceso de revisión se programó y ejecutó para cada una de las versiones liberadas. Luego de cada revisión, se remitían observaciones que eran ajustadas e incorporadas en el desarrollo de una nueva versión.

La Tabla 8 da cuenta de los tipos de errores detectados y reportados en una de las versiones iniciales del aplicativo de precontacto.

Tabla 8. Ejemplo de errores reportados al contratista en el proceso de revisión de software.

ID	Usuario	Folio	Variable	Control Cambios	Tipo	Descripción
1	cap390	102	p1	Error	Validación	Al no encontrar la vivienda (p1 = 2), al cerrar la encuesta aparecen códigos que no corresponden, como 224 o 225.
2	cap390	101	p10a	Modificación	Salto	Al registrar que la persona "no tiene teléfono" aparece alerta de error de validación y se cierra la encuesta. Debería pasar a pregunta p11a y consultar teléfono de otro integrante del hogar.
3	cap390	103	p10a	Modificación	Salto	Al registrar que la persona "rechaza entregar número de teléfono" aparece alerta de error de validación y se cierra la encuesta. Debería pasar a pregunta p11a y consultar teléfono de otro integrante del hogar.
4	cap390	103	p10b/p11b	Modificación	Otros	Especificar en el texto de teléfono y teléfono celular cuantos dígitos se espera ingresar.
5	cap391	101	p1	Error	Validación	Al marcar p1 = 2, al cerrar visita aparece en como ubicación registrada, cuando no se ha logrado ubicar dirección.
6	cap391	102	p11b	Modificación	Validación	Al pasar a p11b "teléfono de otro integrante del hogar", permite ingresar el mismo teléfono que se ingresó anteriormente. Se pueden ingresar 3 veces el mismo número.
7	cap392	102	p12	Error	Otros	Luego de tomar fotografía y aceptar, la app arroja error: "Casen Precontacto 2020 continúa fallando".

Fuente: Elaboración propia

Este proceso de ejecución se implementó a cada una de las 20 versiones de pruebas, antes de llegar a la versión con la que se inició el trabajo de campo. En cada una de estas versiones se reportaron diferentes tipos de incidencias, que se detallan a continuación:

- **Error:** la aplicación no se ajusta a lo solicitado, por ejemplo, cuando no se ha implementado algún tipo de validación requerida.
- **Vigente:** corresponde a errores detectados en una versión previa de la aplicación y reportados nuevamente en la versión en revisión.
- **Modificación:** se solicita hacer un cambio a lo solicitado inicialmente por parte de MDSF.

En la siguiente tabla se detallan cada una de las versiones liberadas y revisadas.

Tabla 9. Versiones software CAPI Casen en Pandemia 2020

Actividad	Estado	Fecha
Pruebas sobre versiones 1.0.1 a 1.0.6	Finalizada	22/09/2020 al 24/09/2020
Pruebas sobre versión 1.0.7	Finalizada	24/09/2020 al 4/10/2020
Pruebas sobre versión 1.0.8	Finalizada	05/10/2020
Pruebas sobre versión 1.0.9	Finalizada	05/10/2020
Pruebas sobre versión 1.0.10	Finalizada	6/10/2020 al 7/10/2020
Pruebas sobre versión 1.0.11	Finalizada	08/10/2020
Pruebas sobre versión 1.0.13	Finalizada	19/10/2020
Pruebas sobre versión 1.0.14	Finalizada	20/10/2020
Pruebas sobre versión 1.0.16	Finalizada	22/10/2020
Pruebas sobre versión 1.0.17	Finalizada	23/10/2020
Pruebas sobre versión 1.0.18	Finalizada	26/10/2020
Pruebas sobre versión 1.0.19	Finalizada	29/10/2020
Pruebas sobre versión 1.0.20	Finalizada	06/11/2020

Fuente: Servicio de Testing Casen en Pandemia 2020, Birchman Group

Cabe destacar que, entre la primera y la última versión entregada, la tasa de resolución de errores fue de 99,9%. Asimismo, los errores restantes no ponían en riesgo la implementación del aplicativo en el trabajo de campo y fueron resueltos oportunamente dentro de la misma versión.

No obstante, durante el desarrollo del trabajo de campo, estas pruebas continuaron ejecutándose por parte tanto del contratista como del equipo del Ministerio de Desarrollo Social y Familia.

Bibliografía

- Caeyers B., Chalmers N., De Weerd J. (2010)** “A Comparison of CAPI and PAPI through a Randomized Field Experiment”. Investigación financiada por el Banco Mundial.
- Caviglia J., Hall S., Mullan K., Macintyre C., Bauch C., Harris D., Sills E., Roberts D., Toomey M., Cha H. (2012)** “Improving Household Surveys Through Computer-Assisted Data Collection: Use of Touch-Screen Laptops in Challenging Environments”. USA.
- Centro de Encuestas y Estudios Longitudinales de la UC (2017)** “Resultados Prueba de Campo Casen 2017 Papel/Tablet” Presentación Panel asesor Casen 2017.
- Centro de Microdatos Departamento de Economía Universidad de Chile (2017)** “Proyecto: Evaluación uso de Tablets en Encuesta Casen” Informe 4.
- Choumert J., Spence C., Taylor C. (2016)** “Using CAPI to Improve and Evaluate the Quality of Socioeconomic Surveys”. High Wycombe, United Kingdom/Bukoba, Tanzania.
- Díaz de Rada, Vidal (2015)** “Calidad de los datos de preguntas de batería en encuestas presenciales: una comparación de un estudio con cuestionario en papel y en formato electrónico”. Revista Española de Investigaciones Sociológicas, 152: 167-178.
- Hunter J., Landreth A. (2005)** “Analyzing Interviewer/Respondent Interactions while Using a Mobile Computer-Assisted Personal Interview Device”, U.S. Census Bureau.
- Laurie H. (2003)** “From PAPI to CAPI: consequences for data quality on the British Household Panel Study”, Working Papers of the Institute for Social and Economic Research, paper 2003-14.
- Martin Jean (1993)** “PAPI to CAPI: The OPC Experience”. Office of Population Censuses and Surveys, London UK.
- Martínez M., Moratilla T. (2000)** “La Captura de Datos Asistida por Ordenador en la Encuesta de Población Activa” Instituto Nacional de Estadística, España.
- Nielsen, J. (2012)** Usability 101: Introduction to Usability. NN Group Recuperado de <https://www.nngroup.com/articles/usability-101-introduction-to-usability>
- Rosero L., Hidalgo J., Antich D., Seligson M. (2005)** “Improving the Quality and Lowering Costs of Household Survey Data Using Personal Digital Assistants (PDAs). An Application for Costa Rica”, Centro Americano de Población de la Universidad de Costa Rica.
- Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2016)**. Designing the user interface: strategies for effective human-computer interaction. Pearson.
- Schraepfer P., Schupp J., Wagner G. (2010)** “Changing from PAPI to CAPI: Introducing CAPI in a Longitudinal Study”, Journal of Official Statistics, Vol. 26, No. 2, pp. 233–269.

Anexo 1: Cuadro comparativo resultados PAPI vs CAPI, Survey CTO (2015)

Ámbito	PAPI	CAPI
Tipo de preguntas	Los formularios pueden permitir respuestas tanto de opción múltiple como de escritura.	Los formularios pueden permitir respuestas de opción múltiple y escrita; grabar audio o video, realizar cálculos, capturar firmas, permitir a los encuestados dibujar en las fotos y más. En caso de encuestas que poseen variadas preguntas abiertas o con escritura, existe un efecto tiempo a considerar.
Hardware	Las computadoras serán necesarias para la entrada de datos.	Se necesitarán teléfonos inteligentes, tabletas, computadoras portátiles o netbooks para que cada uno de los encuestadores realice entrevistas y registre las respuestas. El hardware es uno de los costos clave en la recopilación de datos CAPI.
Software	Será necesario el software de entrada de datos.	Diferentes opciones de software están disponibles. Los precios, la dificultad de programación, el soporte y la flexibilidad varían ampliamente.
Supervisión	Se requieren grandes esfuerzos en supervisión principalmente realizada en terreno para observar que todos los protocolos preestablecidos se cumplan.	Se puede capturar la ubicación de los encuestadores a través de GPS.
Impresión	Los cuestionarios de la encuesta en papel deben estar impresos. Así como tarjeteros y si lo amerita hojas de ruta. La impresión es uno de los costos clave en la recopilación de datos PAPI.	No se requiere impresión de cuestionarios y material de apoyo para una encuesta CAPI.
Capacitación	Los patrones de salto complejos pueden ser difíciles de dominar por los encuestadores, y puede requerir mucha práctica para que los encuestadores ingresen de manera confiable respuestas válidas en todos los campos. Por lo tanto, el período de entrenamiento necesario podría ser largo.	Dado que el dispositivo de recolección de datos maneja los patrones de omisión y la validación de campo, los encuestadores pueden capacitarse más rápidamente. Sin embargo, aquellos que no estén acostumbrados a usar la tecnología pueden necesitar más tiempo.
Transporte y almacenamiento	Se debe diseñar un sistema seguro para llevar las encuestas desde el campo a la oficina y para almacenar las encuestas en la oficina. Una vez que los datos se ingresan en la computadora, se deben almacenar de forma segura (generalmente, mediante el cifrado de datos).	Por lo general, las encuestas se “transportan” a través de una conexión USB, una red local de wifi o Internet, y luego se “almacenan” en un servidor o disco duro local. Deben tomarse las precauciones adecuadas (generalmente mediante el cifrado de datos) para garantizar que los datos estén seguros.

Entrada de datos	Los datos deben ingresarse en la plantilla programada por digitadores entrenados para el ingreso de datos. La entrada de datos es uno de los costos clave en la recopilación de datos PAPI. Esta etapa puede introducir nuevos errores, por concepto de digitación. Si bien existe la doble digitación, este error no se elimina del todo. También podrían darse problemas por lectura óptica, donde cifras con borrones o mal escritas pueden ser mal interpretadas por el lector.	Los datos se digitalizan instantáneamente a medida que se ingresan en la Tablet.
Control de calidad	Por lo general, el equipo de la encuesta incluirá "supervisores", cuyo trabajo consiste en examinar cuidadosamente los cuestionarios completados y detectar los errores del encuestador. El cuestionario puede ser enviado al campo para su corrección.	El flujo de preguntas está automatizado, por lo que el encuestador no necesita decidir qué pregunta viene a continuación. Muchas otras medidas de control de calidad son posibles con CAPI, como las verificaciones lógicas, la información de llenado previo y el establecimiento de restricciones en los rangos de respuesta. Los datos también están disponibles de inmediato en un formato digital. En caso de existir un error en el ingreso de dato del encuestador o incluso en la respuesta del encuestado, este podría ser identificado al momento y ser corregido inmediatamente.
Edición de datos	Una vez que se completa la entrada de datos, puede comenzar el proceso de edición.	El proceso de edición de datos puede comenzar después del primer día de levantamiento. Los encuestadores aún pueden regresar a cualquier hogar si necesitan hacer una corrección.
Disponibilidad de la información	Por lo general, los datos solo están disponibles después de haberlos ingresado y editado. Esto puede ser semanas o meses después de la recopilación de datos.	En la configuración típica donde los equipos de campo cargan de forma segura los datos a un servidor, los datos están disponibles para el día dos de recolección de datos.

Fuente: Survey CTO (2015), <https://www.surveycto.com/best-practices/capi-vs-papi/>