Metaevaluación de las evaluaciones elaboradas dentro de la ANE y otras evaluaciones de programas nacionales en Costa Rica

Encargada por: FOCEVAL

Realizada por: Juan Murciano

Fecha de encargo: 26 de octubre de 2016

Fecha de entrega: 19 de diciembre de 2016

Fecha de actualización 20 de marzo de 2017

Índice de contenidos

Lista de acrónimos **3**

Lista de nombres abreviados de las evaluaciones metaevaluadas **3**

1. Introducción **1**

1.1. Delimitación conceptual de la metaevaluación, utilidad y vínculo con los estándares internacionales de evaluación 1

1.2. Descripción del objeto de la metaevaluación 2

1.3. Descripción de la metodología aplicada para la metaevaluación 6

2. Metaevaluación de las evaluaciones **11**

2.1. Síntesis de las evaluaciones evaluadas 11

2.1.1 Evaluación 01: DRAT 12
2.1.2 Evaluación 02: EDU-INCLUSIVA 14
2.1.3 Evaluación 03: GERMINADORA 16
2.1.4 Evaluación 04: PND 18
2.1.5 Evaluación 05: REDCUDI 21
2.1.6 Evaluación 06: CEN-CINAI 23
2.1.7 Evaluación 07: EMPLÉATE 25
2.1.8 Evaluación 08: OAPVD 27
2.1.9 Evaluación 09: IMAS 29
2.1.10 Evaluación 010: PRENATAL 31
2.1.10 Evaluación 011: PREVENCION CÁNCER (PyA-Cáncer) 33

2.2. Hallazgos: apreciación general de las evaluaciones 35

2.2.1. Sinopsis de la calificación de las evaluaciones analizadas por la metaevaluación. 35
2.2.1. Descripción de patrones y tendencias generales 36

3. Conclusiones **43**

3.1. Valoración del indicador de FOCEVAL 43

3.2. Factores determinantes de la calidad general de las evaluaciones. 45

4. Recomendaciones derivadas de la metaevaluación **48**

4.1. Calidad y utilidad de las evaluaciones y los procesos evaluativos 48

4.2. Calidad de los Términos de Referencia 50

4.3. Calidad y utilidad de la metaevaluación 51

5. Anexos **53**

5.1 Términos de referencia de la metaevaluación 53

5.2. Documentos utilizados 61

5.3 Tabla sintética de valoración cuantitativa de informes de evaluación y TdR 62

5.4 Ficha de registro de evaluaciones de la metaevaluación 66

4.4.1 Evaluación 01: DRAT 67
4.4.2 Evaluación 02: EDU-INCLUSIVA 76
4.4.3 Evaluación 03: GERMINADORA 89
4.4.4 Evaluación 04: PND 100
4.4.5 Evaluación 05: REDCUDI 111
4.4.6 Evaluación 06: CEN-CINAI 123

4.4.7 Evaluación 07: EMPLÉATE 135
4.4.8 Evaluación 08: OAPVD 148
4.4.9 Evaluación 09: IMAS 161
4.4.10 Evaluación 010: PRENATAL 170
4.4.11 Evaluación 011: PyA-Cáncer 180

Índice de tablas y gráficos

Tabla 1. Objeto de metaevaluación: evaluaciones según modalidad de apoyo y ejecución. 3

Tabla 1. Modalidades de apoyo FOCEVAL a las evaluaciones. 3

Tabla 2. Evaluaciones según año de realización, tipo y modalidad de apoyo FOCEVAL 5

Tabla 3. Ponderación de criterios y dimensiones en la valoración cuantitativa 8

Tabla 4. Criterios y preguntas de metaevaluación 9

Tabla 5. Sinopsis de calificaciones de las evaluaciones según criterios aplicados y dimensiones 35

Tabla 6. Puntuaciones de las evaluaciones en el criterio de Imparcialidad e independencia 36

Tabla 7. Puntuaciones de las evaluaciones en el criterio de Equidad, participación y transparencia . 37

Tabla 8. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Metas, objetivos y modelo de cambio 39

Tabla 9. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Diseño y métodos 39

Tabla 10. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Procedimiento de selección 40

Tabla 11. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Análisis de datos 41

Tabla 12. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Criterios de evaluación 41

Tabla 13. Puntuaciones de las evaluaciones en el criterio de Calidad del informe y utilidad 42

Tabla 14. Puntuaciones de las evaluaciones en el criterio de Calidad del informe y utilidad 42

Tabla 15. Grado de cumplimiento de las evaluaciones en los cinco criterios de metaevaluación (alto-medio-bajo) 44

Tabla 16. Grado de cumplimiento de las evaluaciones en los cinco criterios de metaevaluación (alto-medio-bajo) 45

Gráfico 1. Puntuaciones de las evaluaciones según criterios aplicados 36

Gráfico 2. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud según dimensiones 38

Lista de acrónimos

ANE	Agenda Nacional de Evaluaciones
BID	Banco Interamericano de Desarrollo
BMZ	Ministerium für wirtschaftliche Zusammearbeit und Entwicklung (Ministerio Federal de Cooperación Económica y Desarrollo, Alemania)
CCSS	Caja Cistarricense del Seguro Social
DEval	Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (Instituto Alemán de Evaluación de la Cooperación al Desarrollo)
ECD	Evaluation capacity development
FOCEVAL	Programa de Fomento de Capacidades en Evaluación en diversos países de América Latina
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH
MIDEPLAN	Ministerio de Planificación Nacional y Política Económica
MEP	Ministerio de Educación Pública .
TdR	Términos de Referencia de la evaluación/metaevaluación

Lista de nombres abreviados de las evaluaciones metaevaluadas

DRAT	Evaluación de Impacto del "Distrito de Riego Arenal-Tempisque", de Servicio Nacional de Áreas Subterráneas, Riego y Avenamiento (SENARA).
EDU-INCL	Evaluación del proyecto "Apoyo a la Gestión pedagógica a Centros educativos de calidad con orientación inclusiva", del Ministerio de Educación Pública (MEP).
Germinadora	Evaluación del proyecto interinstitucional "Germinadora de Desarrollo Organizacional Empresarial, Asociativo y Comunitario".
PND	Evaluación del "Diseño Metodológico y Proceso de Elaboración del PND 2011 – 2014", de MIDEPLAN.
REDCUDI	Evaluación del programa interinstitucional "Red Nacional de Cuido y Desarrollo Infantil (REDCUDI)".
CEN-CINAI	Evaluación del programa "Centros de Educación y Nutrición, y Centros Infantiles de Atención Integral (CEN-CINAI)", del Ministerio de Salud.
EMPLÉATE	Evaluación del programa "Empléate", del Ministerio de Trabajo y Seguridad Social.
OAPVD	Evaluación del "Programa de Atención de la Oficina de Atención y Protección a la Víctima de Delito", del Poder Judicial de Costa Rica.
IMAS	Evaluación de efectos generados por los programas sociales que ejecuta el IMAS, 2009
PRENATAL	Evaluación del programa de Atención Prenatal en el Primer Nivel de Atención del Ministerio de Salud
PyA-Cáncer	Evaluación participativa de los servicios de prevención y atención del cáncer en Valle la Estrella, provincia de Limón, (Costa Rica, 2017)

1. Introducción

FOCEVAL (Fomento de Capacidades en Evaluación en Costa Rica y países seleccionados de América Latina) es un proyecto de fortalecimiento de capacidades en evaluación financiado con fondos del BMZ implementado por DEval, la contraparte principal del proyecto es el Ministerio de Planificación Nacional y Política Económica de Costa Rica (MIDEPLAN).

El proyecto FOCEVAL fue concebido como una plataforma interinstitucional para apoyar la creación de capacidades en evaluación tanto en Costa Rica como en otros países de América Latina. Su objetivo es fortalecer el rol que juega la evaluación en Costa Rica y en países seleccionados de América Latina en la toma de decisiones y en la conducción política.

Para lograr este objetivo, el proyecto tiene tres componentes: capacitación, regionalización e institucionalización. En este último se enmarca como una actividad principal el acompañamiento de la ejecución de la Agenda Nacional de Evaluación (ANE) en Costa Rica. El órgano responsable del diseño e implementación de la ANE es MIDEPLAN, institución rectora del Sistema Nacional de Evaluación de Costa Rica.

FOCEVAL apoya la implementación de varias evaluaciones de la ANE. Para aprender cómo mejorar la calidad de las evaluaciones, el proceso evaluativo y el uso de las evaluaciones (y la interacción entre estas dimensiones), se quiere celebrar dos talleres a principios del 2017. Como un insumo importante para estos talleres se requiere hacer una metaevaluación de por lo menos 10 evaluaciones realizadas en Costa Rica en los últimos años.

Es en este contexto en el que surge la necesidad de realizar esta metaevaluación de las evaluaciones elaboradas dentro de la ANE y otras evaluaciones de programas nacionales en Costa Rica, incluyendo los respectivos términos de referencia.

1.1. Delimitación conceptual de la metaevaluación, utilidad y vínculo con los estándares internacionales de evaluación

Concepto de metaevaluación. Conocer la utilidad de la evaluación, su rigor metodológico, el acierto en el manejo de recursos y enfoques o el desempeño profesional y ético del equipo evaluador son preocupaciones que nos sitúan ante la metaevaluación, o lo que es lo mismo, ante la *evaluación de la evaluación y su corolario la evaluación de los evaluadores*, conceptos introducidos por Scriven en 1968 (Murciano, 2007:175). Según Stufflebeam (1987: 52), el objetivo último de la metaevaluación es asegurar la calidad de los servicios evaluativos, evitar o enfrentarse a prácticas ilegales o que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover mayor comprensión de la empresa evaluativa.

Bustelo (2001; 196) le asigna dos significados a la metaevaluación; uno centrado en la "descripción, análisis y valoración de los procesos de evaluación" y otro asociado a "su función como control de calidad de las evaluaciones".

> "Para Schwandt y Halpern (1988: 14), la metaevaluación se inscribiría en una serie de aproximaciones sistemáticas para el seguimiento, fomento y comprobación de la calidad de un servicio profesional como es, en este caso, la evaluación. El interés por la metaevaluación surge, en cierto modo, cuando se da una preocupación por la calidad de las evaluaciones y la necesidad de control de las mismas. En realidad tiene que ver con lo que Scriven define como la característica autoreferenciada o reflexiva que debe tener la evaluación, es decir, con la idea de "la evaluación empieza en casa" y con la de "¿quién evalúa al evaluador?" (Scriven, 1993: 51-52)". Citado en Bustelo (2001; 196)

Esta metaevaluación toma ambos significados descritos por Bustelo (2001), con un alto énfasis en el control de calidad en línea con la aproximación de Schwandt y Halpern (1988), en coherencia con la finalidad última de FOCEVAL.

Utilidad y uso de la metaevaluación. *Su verdadero sentido deriva de su orientación práctica*, dado que su uso debe servir no sólo para tomar mejores decisiones sobre políticas y programas públicos, sino para propiciar el aprendizaje de quienes evalúan y de quienes encargan evaluaciones, contribuir al perfeccionamiento de la profesión y, en última instancia, fortalecer la transparencia y rendición de cuentas a los ciudadanos (Osuna et. al., 2012). Es justo en este punto, en el que resulta paradójico que su "mayor" desarrollo se haya producido

en el ámbito académico y no en el seno de la Administración Pública. La mayoría de referencias se circunscriben a tesis doctorales, investigaciones o artículos científicos y no a trabajos demandados por Unidades de Evaluación. Sin cuestionar la indudable aportación de los trabajos académicos, su utilidad para inducir cambios efectivos en la mejora de la práctica evaluadora es limitada. La consolidación de Unidades de evaluación está propiciando una "nueva oleada de metaevaluaciones" demandadas por estos organismos que atesoran ya una trayectoria relevante de evaluaciones y que tienen capacidad para orientar la práctica evaluativa mejorando sus políticas de evaluación.

Es aquí donde se ubica esta metaevaluación encargada por FOCEVAL, y por tanto su relevancia y su carácter innovador, pues es el propio programa FOCEVAL el que tras una trayectoria relevante de apoyo y promoción de la evaluación de políticas públicas en Costa Rica, encarga esta metaevaluación para contribuir a la mejora de la calidad de los procesos. Además, contar con la participación directa de MIDEPLAN, como institucionalidad pública que ostenta la rectoría del Sistema Nacional de Evaluación, ofrece una importante capacidad para inducir cambios en la práctica evaluativa pues, mediante el lidrazgo en la ANE, puede modelar sus encargos y modalidades de apoyo para optimizar la mejora de calidad de las evaluaciones.

Estándares internacionales. Las propuestas para hacer metaevaluaciones tratan de acotar el deber ser de una evaluación configurando criterios o principios profesionales, priorizando lo técnico, lo político o lo ético, u optando por algún enfoque particular (García, 2011:35). Aportaciones más recientes como la vía comprensiva de Stake se apoyan en las percepciones cualitativas de quienes participaron en el proceso y quienes metaevalúan (Stake, 2006: 256). Entre todas, han sido las propuestas apoyadas en códigos deontológicos de sociedades, grupos de evaluación y organismos internacionales las que han tenido más eco, destacando las del Joint Committee on Standards for Educational Evaluation (1980) o las de Naciones Unidas. En lo más reciente cabe destacar el proyecto de *Formulación de estándares de evaluación para América Latina y el Caribe* desarrollado por FOCEVAL y la ReLAC bajo el auspicio de DEval, pues aúna una exhaustiva revisión bibliográfica, una consulta a la comunidad evaluadora y aportaciones de expertos. Su propuesta abarca tres dominios o ámbitos: teórico-metodológico; ético; y comunicativo[1].

Esta metaevaluación toma como punto de partida la lista de criterios y preguntas predefinida por FOCEVAL, la cual trata de sintetizar los estándares internacionales de evaluación, focalizando su atención en los informes de evaluación y en sus TdR, para su realización a partir del análisis documental, por una persona experta en la materia.

1.2. Descripción del objeto de la metaevaluación

Esta metaevaluación toma como objeto un conjunto de once evaluaciones de programas nacionales de Costa Rica, las cuales se presentan en la Tabla 1. Para delimitar este objeto evaluativo, se puede atender en un primer momento a dos variables: por un lado el momento temporal en el que se realizan y, por otro, la modalidad de conformación de los equipos de evaluación y de gestión de la evaluación.

Por un lado, si se atiende al momento temporal en el que se realiza la evaluación respecto al proyecto FOCEVAL, las evaluaciones objeto de metaevaluación se pueden clasificar como:

1. Evaluaciones bajo la primera etapa del proyecto FOCEVAL. Estas evaluaciones se denominan evaluaciones piloto y se corresponden con evaluaciones que fueron presentadas y seleccionadas en el Concurso convocado en noviembre – diciembre 2012 por el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) y la GIZ (Agencia Alemana de Cooperación Alemana), aprovechando los alcances del proyecto FOCEVAL.
2. Evaluaciones bajo la segunda etapa del proyecto FOCEVAL. Estas evaluaciones forman parte de la Agenda Nacional de Evaluaciones (ANE), y fueron por tanto presentadas por los órganos responsables de los programas y proyectos para su inclusión en la ANE según el protocolo establecido al efecto desde MIDEPLAN.
3. Otras evaluaciones. Estas evaluaciones no tienen por qué haberse desarrollado bajo el amparo del proyecto FOCEVAL y su gestión y realización corresponde a determinados ministerios. En este caso se han seleccionado evaluaciones del Ministerio de Salud así como la primera evaluación participativa apoyada.

[1] Documento de trabajo interno Taller Estándares de Evaluación para América Latina. IV Conferencia ReLAC, 2015

Tabla 1. Objeto de metaevaluación: evaluaciones según modalidad de apoyo y ejecución.

Evaluaciones piloto [FOCEVAL I]	Evaluaciones ANE [FOCEVAL II]	Evaluaciones académicas de aprendizaje
[Consultores nacionales e internacionales] **MODALIDAD CONSULTORÍA EXTERNA**	[Evaluadores nacionales y gestionadas por un "equipo gestor" con asesoría de consultores internacionales] **MODALIDAD PADRINAZGO**	[Evaluadores jóvenes y emergentes de la Maestría en Evaluación (UCR) con apoyo de FOCEVAL (nº8) y BID (nº9)] **MODALIDAD EVALUADORES EMERGENTES**
1. Evaluación de Impacto del "**D**istrito de **R**iego **A**renal-**T**empisque", de Servicio Nacional de Áreas Subterráneas, Riego y Avenamiento (SENARA). **DRAT**	**5.** Evaluación del programa interinstitucional "Red Nacional de Cuido y Desarrollo Infantil (**REDCUDI**)".	**8.** Evaluación del "Programa de Atención de la **O**ficina de **A**tención y **P**rotección a la **V**íctima de **D**elito", del Poder Judicial de Costa Rica. **OAPVD**
2. Evaluación del proyecto "Apoyo a la Gestión pedagógica a Centros educativos de calidad con orientación inclusiva", del Ministerio de Educación Pública (MEP). **EDU-INCL**	**6.** Evaluación del programa "Centros de Educación y Nutrición, y Centros Infantiles de Atención Integral (**CEN-CINAI**)", del Ministerio de Salud.	**9.** Evaluación del programa "**Avancemos**" del Ministerio de Educación Pública, junto con otras instituciones. **IMAS** **Evaluación Ministerio de Salud**
3. Evaluación del proyecto interinstitucional "**Germinadora** de Desarrollo Organizacional Empresarial, Asociativo y Comunitario".	["Equipo gestor" evaluador con asesoría de consultoras nacional e internacional] **MODALIDAD APRENDER-HACIENDO**	[Equipo Evaluador Interinstitucional, coordinado por la Unidad de Evaluación de Impacto de las Acciones en Salud]
4. Evaluación del "Diseño Metodológico y Proceso de Elaboración del **PND** 2011 – 2014", de MIDEPLAN.	**7.** Evaluación del programa "**Empléate**", del Ministerio de Trabajo y Seguridad Social.	**10.** Evaluación de la Calidad de la Atención **Prenatal** en el Primer Nivel de Atención, Región Brunca y Región Huetar Caribe (Costa Rica, 2012). **Evaluaciones participativas** [Equipo evaluador (interno y con stakeholders) con asesorías especializadas y bajo la coordinación de DEval-Foceval] **11.** Evaluación participativa de los servicios de prevención y atención del cáncer en Valle la Estrella, provincia de Limón, (Costa Rica, 2017)

Fuente: TdR de la metaevaluación

Por otro lado, se puede atender a la modalidad en la que se diseña, ejecuta y gestiona el proceso evaluativo. Desde esta óptica, las evaluaciones conforman procesos en los que desde el proyecto FOCEVAL se han experimentado diferentes formas de fomentar el desarrollo de capacidades en evaluación, desplegando diversas modalidades de apoyo a los procesos de evaluación, en las que se encuentran implicadas tanto las organizaciones responsables de los programas evaluados, MIDEPLAN como rector del Sistema Nacional de Evaluación, y la conformación de los equipos evaluadores.

Tabla 1. Modalidades de apoyo FOCEVAL a las evaluaciones.

MODALIDADES DE APOYO FOCEVAL	EQUIPO EVALUADOR		EQUIPO GESTOR DE LA EVALUACIÓN
	[Contratación de Expertos/as Nacionales/Internacionales para realizar la evaluación]		**MODALIDAD PADRINAZGO** **[Experto/a Internacional apoya al Equipo Gestor]**
	MODALIDAD APRENDER HACIENDO **[Experto/a Internacional/Nacional apoya al Equipo Gestor=Equipo Evaluador]**		
	MODALIDAD EVALUACIÓN ACADÉMICA **[Contratación de Evaluadores jóvenes/emergentes (Maestria de Evaluación de Políticas Públicas–UCR) para realizar la evaluación]**		**Equipo Gestor de la evaluación**
	MODALIDAD EVALUACIÓN EXTERNA **[Contratación de Expertos/as Nacionales/Internacionales para realizar la evaluación]**		**Equipo Gestor de la evaluación**
	MODALIDAD ASISTENCIA TÉCNICA **[Asesoramiento de FOCEVAL –GIZ y MIDEPLAN al Equipo evaluador interno]**		
	Submodalidad Evaluación mixta	**Submodalidad Evaluación participativa**	

Fuente: elaboración propia a partir de los TdR de la metaevaluación e Informes finales de evaluación.

Bajo este prisma, se encuentran las siguientes modalidades de apoyo y evaluaciones apoyadas:

1. Modalidad Aprender-Haciendo. Bajo esta modalidad el equipo evaluador es el propio equipo gestor de la evaluación, reuniendo por tanto integrantes tanto de MIDEPLAN como rector del sistema nacional de

evaluación, como de la institución responsable de la implementación del proyecto/programa que se evalúa. Este equipo evaluador/gestor recibe el acompañamiento de consultores/as para el diseño y realización de la evaluación.

En la primera fase del proyecto FOCEVAL se desarrolló bajo esta modalidad la evaluación del Proyecto Apoyo a la Gestión pedagógica de centros educativos de calidad con orientación inclusiva (**EDU-INCL**), evaluación presentada y seleccionada en el Concurso de evaluaciones piloto que, durante los meses de noviembre y diciembre 2012 fue convocado por el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) y la Agencia Alemana de Cooperación (GIZ). Se conformó para tales efectos un grupo de trabajo con la participación de representantes del MIDEPLAN (con funcionarios de la Unidad de Acciones Estratégicas (UAE) del Área Evaluación y Seguimiento (AES)), así como representantes de GIZ y del Ministerio de Educación Pública. Este grupo fue asesorado por una experta nacional en evaluación y en la política evaluada (educativa). Asimismo, se apoyó con esta modalidad durante la primera etapa de FOCEVAL la Evaluación del proyecto interinstitucional "Germinadora de Desarrollo Organizacional Empresarial, Asociativo y Comunitario" (**Germinadora**). Estas dos evaluaciones (EDU-INCL y Germinadora) contaron con acompañamiento internacional.

En la segunda fase del proyecto FOCEVAL en el marco de la ANE, se desarrolló bajo esta modalidad la evaluación del programa **Empléate**, recibiendo el equipo gestor/evaluador el acompañamiento de dos consultoras, una nacional y otra internacional, expertas en evaluación y en la política evaluada (empleo y formación). Cada etapa de la evaluación estuvo precedida de un taller de capacitación por parte de las consultoras que los integrantes del equipo evaluador fueron poniendo en práctica durante el diseño, programación e implementación de la evaluación.

2. Modalidad Padrinazgo. Esta modalidad corresponde a evaluaciones ejecutadas por evaluadores (nacionales) y gestionadas por un "equipo gestor" compuesto por funcionarios y funcionarias de las instituciones involucradas en los programas y de MIDEPLAN como órgano rector del Sistema Nacional de Evaluación. La evaluación la realizan equipos expertos nacionales y los equipos gestores de las evaluaciones son asesorados por consultores internacionales (padrinos) que seleccionó FOCEVAL y los puso a disposición del proceso evaluativo–según los TdR de la metaevaluación-, concibiéndose originalmente su apoyo -según los TdR de la evaluación de REDCUDI- tanto al equipo gestor como al equipo evaluador, si bien quedó circunscrito el apoyo al equipo gestor de la evaluación.

   Dos evaluaciones, realizadas en la segunda fase del proyecto FOCEVAL en el marco de la ANE, fueron apoyadas bajo esta fórmula: las evaluaciones de los programas **CEN-CINAI** y **REDCUDI**.

3. Modalidad Evaluación Académica [Evaluadores jóvenes y emergentes]. Esta modalidad consiste en desplegar evaluaciones académicas de aprendizaje, en las que el equipo evaluador está conformado por evaluadores jóvenes y emergentes de la Maestría de Evaluación de la Universidad de Costa Rica con el apoyo de un órgano externo especializado, en un caso FOCEVAL y en otro el BID. Son evaluaciones en las que los egresados pueden poner en práctica los conocimientos adquiridos y tiene un importante contenido académico.

   Bajo esta modalidad se desplegaron las evaluaciones del "Programa de Atención de la Oficina de Atención y Protección a la Víctima de Delito", del Poder Judicial de Costa Rica (**OAPVD**) y la Evaluación de los efectos de los programas sociales del IMAS (Evaluación del programa "**Avancemos**" del Ministerio de Educación Pública, junto con otras instituciones).

4. Evaluaciones externas. Esta modalidad de apoyo consiste en la contratación de un equipo externo experto en evaluación, nacional o internacional, para que realice de manera independiente la evaluación.

   Bajo esta modalidad se apoyaron durante la primera etapa de FOCEVAL las evaluaciones siguientes: Evaluación de Impacto del "Distrito de Riego Arenal-Tempisque" (**DRAT**), del Servicio Nacional de Áreas Subterráneas, Riego y Avenamiento (SENARA); y la Evaluación del "Diseño Metodológico y Proceso de Elaboración del PND 2011 – 2014" (**PND**), de MIDEPLAN.

5. Otras experiencias: evaluaciones con asistencia técnica al equipo evaluador interno. También se integran en la metaevaluación otras evaluaciones, como la realizada por el Ministerio de Salud sobre el programa de Atención Prenatal en el Primer Nivel de Atención (**PRENATAL**) o la evaluación participativa de los servicios de prevención y atención del cáncer en Valle la Estrella (Prevención Cáncer).

a. Evaluaciones mixtas. En PRENATAL, la evaluación liderada por el Ministerio de Salud, el apoyo de FOCEVAL se concretó en una asistencia técnica centrada especialmente en su etapa de diseño y programación así como en la participación de MIDEPLAN en el Comité revisor, en la elaboración del diseño de la evaluación, en la aplicación de instrumentos, en el proceso de análisis de información y en la elaboración del informe .

b. Evaluaciones participativas. Esta modalidad se concreta en apoyar iniciativas en las que los propios stakeholders demandan la evaluación y conforman el equipo evaluador, ofreciendo FOCEVAL asesorías especializadas, contando con el apoyo de MIDEPLAN y asumiendo DEval la coordinación del mismo[2]. Bajo esta modalidad la primera experiencia apoyada ha sido la evaluación participativa, con la que se apoyó la evaluación de los Servicios de prevención y atención del cáncer en Valle la Estrella, provincia de Limón, (Costa Rica, 2017) (PyA-Cáncer). En esta evaluación el equipo evaluador (Junta de Salud y CCSS) recibió asesoramiento experto (en evaluación participativa y en facilitación de procesos participativos), contó con el apoyo de MIDEPLAN y la Defensoría de Habitantes, así como con la coordinación de FOCEVAL.

Estas dos maneras de delimitar el objeto evaluativo (según etapas de FOCEVAL y según modalidades de apoyo), son relevantes para esta metaevaluación, pues interesa conocer tanto la evolución temporal de la calidad de las evaluaciones desplegadas y apoyadas desde FOCEVAL, como los resultados que se han derivado de las diferentes modalidades de apoyo, su utilidad para desarrollar capacidades de evaluación, su mejora y su capacidad de su replicabilidad.

Tabla 2. Evaluaciones según año de realización, tipo y modalidad de apoyo FOCEVAL

EVALUACION				APOYO FOCEVAL (Modalidad: Cómo y Quién evalúa)			
Código / Nombre		Año	Tipo	Modalidad	Tipo/figura de apoyo (Cómo)	Evaluador (Quién)	Etapa
1	DRAT	2014	Impacto	Externa	Asesoría técnica (al equipo de gestión)	Consultor Nacional	FOCEVAL I
2	EDU-INCL	2014	Procesos, implementación	Aprender haciendo	Asesoría técnica (al equipo de gestión=equipo evaluador)	Consultor Nacional	
3	Germinadora	2014	Evaluacion Intermedia			Consultor Nacional	
4	PND	2013	Diseño	Externa		Consultor Internacional	
5	REDCUDI	2015[3]	Diseño y procesos	Externa con Padrinos	Asesoría técnica (al equipo de gestión)	Consultor Internacional	FOCEVAL II
6	CENCINAI	2015	Procesos			Consultor Internacional	
7	EMPLEATE	2016	Diseño-Procesos-Resultados	Aprender haciendo	Asesoría técnica (al equipo de gestión=equipo evaluador)	Consultor Internacional y Consultor Nacional	
8	OAPVD	2015	Calidad del servicio	Evaluación acadmécica		MEPPD-CEval	
EVALUACION				APOYO FOCEVAL			
9	IMAS (Avancemos)	2011	Resultados (Efectos)	Evaluación acadmécica		MEPPD-BID	FOCEVAL I
10	PRENATAL	2012	Calidad del servicio	Asistencia técnica Evaluación mixta	Asesoría técnica al equipo evaluador (MIDEPLAN-GIZ)	Evaluación interna Ministerio de Salud	
11	PREVENCION Y ATENCIÓN CANCER	2016	Participativa	Asistencia Técnica Evaluación participativa	Asesoría técnica al equipo evaluador (MIDEPLAN-DEval)	Evaluación interna (Juntas de Salud-CCSS)	Otras

Fuente: TdR de la metaevaluación e Informes finales de evaluación.

Desde un enfoque sectorial, las evaluaciones presentan un amplio abanico de sectores a los que principalmente se adscriben los programas evaluados, siendo por tanto también variados los ministerios y organismos responsables de su ejecución. Además de la centrada en el Poder Judicial como uno de los Poderes de la República, las evaluaciones se centraron en los siguientes sectores: Salud (PRENATAL y PyA-Cáncer); Trabajo y Seguridad Social (Empléate); Desarrollo Humano e inclusión social/Infancia (REDCUDI, Avancemos); Sector Productivo (DRAT); Salud, nutrición y deporte/Infancia (CEN-CINAI); Bienestar Social y Familia (Germinadora); Ordenamiento Territorial y Vivienda (PND); y Educación (EDU-INCL).

[2] Esta línea se inicia a partir de la propuesta del Grupo de Coordinación de FOCEVAL de vincular evaluación y participación ciudadana (junio de 2014). Esta línea se inicia con un Encuentro con organizaciones sociales para conversar sobre evaluación (julio de 2015). Continúa con una capacitación en evaluación participativa dirigida a estas organizaciones (noviembre de 2015). Y, deriva, en enero de 2016 en la convocatoria de evalauciones participativas dirigidas a estas organizaciones sociales ofreciéndoles apoyo para elaborar sus propuestas. En esta convocatoria fue seleccionada la propuesta presentada por Consejo Regional de Juntas de Salud de la Región Huétar Atlántico, dando inicio en junio de 2016 la evaluación participativa de los servicios de prevención y atención del cáncer en Valle de la Estrella.

[3] La evalaución de REDCUDI se realizó durante 2015, si bien su informe final se entregó en junio de 2016.

Por otro lado, otra lectura que puede hacerse de las evaluaciones es según el momento en que fueron realizadas, a partir del año de realización y/o entrega del informe final. En este sentido, atendiendo a la fecha de entrega del informe final se cuenta con[2]:

2011	2012	2013	2014	2015	2016	2017
(1) IMAS (Avancemos)	**(1)** PRENATAL	**(1)** PND	**(3)** DRAT, EDUC-INCL, Germinadora	**(3)** REDCUDI, CEN-CINAI, OAPVD	**(2)** EMPLÉATE	**(1)** Prevencion/atención PyA-Cáncer

1.3. Descripción de la metodología aplicada para la metaevaluación

La metaevaluación debe cumplir con tres objetivos estipulados en los TdR que la regulan:

a) Valorar el indicador de FOCEVAL "Las evaluaciones realizadas de forma independiente por las contrapartes del proyecto, dentro del marco de medidas de ECD, cumplen con los criterios de calidad internacionales."

b) Producir resultados que contribuyan a la mejora de los procesos de evaluación en Costa Rica (de MIDEPLAN y otros actores).

c) Elaborar resultados y recomendaciones que sirvan como insumo para dos talleres en los que se elaborarán lecciones aprendidas sobre cómo mejorar la calidad de las evaluaciones, el proceso evaluativo, el uso de las evaluaciones y la interacción entre estos puntos.

Para ello, según indicaban los TdR "la persona experta contratada deberá valorar la calidad metodológica, la validez de los resultados y los términos de referencia de un mínimo de 10 evaluaciones según 30 a 40 criterios pre-establecidos, que corresponden a estándares internacionales y que serán compartidos con la/el experta/o después de haber firmado el contrato". Cabe apuntar que este encargo inicial tuvo una ampliación para incorporar la evaluación de los servicios de prevención y atención del cáncer, cuyo informe esta previsto terminar en febrero de 2017.

La metodología de la metaevaluación toma como punto de partida los objetivos descritos en los TdR y la lista de criterios y preguntas confeccionada por FOCEVAL y remitida para iniciar esta consultoría. Esta lista, según los TdR, concentra y sintetiza las principales directrices internacionales sobre estándares de evaluación y se expone en la página siguiente (Ver Tabla 3).

Esta lista se articula en 5 criterios de evaluación[4]: (1) Imparcialidad e independencia; (2) equidad, participación y transparencia; (3) credibilidad y exactitud; (4) calidad del informe y utilidad; (5) exhaustividad y adecuación de la información). El criterio de Credibilidad y exactitud viene, a su vez, desagregado en 5 dimensiones: (3.1) metas, objeto y modelo de cambio; (3.2) diseño y métodos; (3.3) procedimiento de selección; (3.4) análisis de datos; y (3.5) criterios de evaluación. Cada uno de estos criterios y dimensiones contiene una serie de preguntas sobre sus principales rasgos definitorios.

Para facilitar la respuesta sistemática a cada pregunta formulada se procedió a establecer categorías de respuestas que permitiesen otorgar una calificación, tanto en texto como de forma numérica. Las respuestas en texto especifican en qué medida la evaluación que se metaevalúa contempla todos los atributos a los que se refiere la pregunta, sólo algunos de ellos, o ninguno, o bien que la pregunta no resulta de aplicación. Por su parte, la calificación numérica otorga un valor a cada respuesta, siguiendo una escala de 4 valores (0, 1, 5, 10) según se especifica en el cuadro siguiente:

CATEGORIAS DE RESPUESTA A LAS PREGUNTAS QUE CONTIENE LA LISTA DE CRITERIOS	
VALOR	**DESCRIPCIÓN**

[4] El orden en el que se exponen y aplican los criterios y dimensiones corresponde al otorgado en la Lista confeccionada por FOCEVAL.

Valor "0" (No Aplica)	Esta categoría señala que la pregunta *no puede aplicarse* a la evaluación, por lo que no puede sumar en la puntuación total de la evaluación, ni se contempla este ítem a lo hora de establecer en cuantos ítems se desagrega el criterio o subcriterio de metaevaluación para obtener su puntuación promedio, de forma que no se vea penalizada por un ítem del que no puede dar respuesta. Se puede decir, que esta categoría es la que permite una adaptación del esquema de valoración al contexto particular de cada proceso evaluativo que se juzga. (Este valor se muestra en color **negro** en las tablas y en las Fichas de registro de cada evaluación).
Valor "1"	Indica una respuesta *negativa* a la pregunta que se formula. (Este valor se muestra en color **rojo** en las tablas y en las Fichas de registro de cada evaluación).
Valor "5"	Indica una respuesta *intermedia* a la pregunta que se formula. Normalmente si la pregunta presenta varias implicaciones para la evaluación, esta pregunta se concreta en que sólo se aprecian algunos de ellos, o si refieren a una graduación que la calificación es intermedia o moderada. (Este valor se muestra en color **naranja** en las tablas y en las Fichas de registro de cada evaluación).
Valor "10"	Indica una respuesta *positiva* a la pregunta que se formula (Este valor se muestra en color **verde** en las tablas y en las Fichas de registro de cada evaluación).

La elección de esta escala se fundamenta en que muchas de las preguntas formuladas casi precisan una respuesta dicotómica, Sí/No, en la que sólo es factible establecer una respuesta intermedia. Asimismo, dado que algunas de las preguntas, especialmente las relativas al grado de respuesta que ofrecen las evaluaciones a determinados criterios de evaluación, no puede asegurarse que sean de aplicación a todos los casos[5], se hizo necesario incorporar una categoría para reflejar esta situación ("No Aplica"). Junto a ello, se optó por contemplar como complemento a cada respuesta un apartado de observaciones cualitativas, en el que se tratan de sistematizar tanto las evidencias que sustentan la calificación otorgada, como aquellas otras apreciaciones que resultan de interés para aspecto que se juzga en cada caso.

En los anexos se presentan las categorías de respuesta (4.2) y la Ficha de registro de cada evaluación (4.4). En la Tabla 4 que se presenta más adelante, se expone este conjunto de criterios y preguntas de metaevaluación.

Una vez establecidas las categorías de respuesta, se definió el **método de cálculo** para obtener una puntuación numérica resultante de la aplicación de los criterios y preguntas a cada evaluación.

- Cada pregunta obtiene una puntuación según la categoría de respuesta seleccionada (0,1, 5 o 10).
- Para obtener la calificación de cada criterio o dimensión se aplica el promedio de las puntuaciones obtenidas por el conjunto de preguntas que los componen. Es decir, se suman las puntuaciones de todas las preguntas que conforman un criterio/dimensión y se divide por el número de preguntas que aplican a dicho criterio/dimensión. Esto último es así, ya que en caso de que haya alguna pregunta en la que la calificación sea 0 ("No aplica"), esa pregunta no se tiene cuenta en el denominador del promedio, para no penalizar a la evaluación que se califica por una pregunta de la que no puede dar respuesta (bajaría sobremanera su promedio), otorgando mayor flexibilidad a la calificación al contemplar la diversidad de contextos evaluativos.
- El resultado final se obtiene estableciendo ponderaciones sobre los criterios y dimensiones.
  - El supuesto de partida fue otorgar igual ponderación a todos los criterios y dimensiones. De lo que se deduce que si se otorgan 10 puntos a cada criterio, el criterio de Credibilidad y exactitud, al estar desagregado en 5 dimensiones, obtiene una puntuación de 50 puntos sobre el total; recibiendo 10 puntos el resto de criterios. Lo que daría una puntuación total máxima de 90 puntos para una evaluación.
  - Se acordó con DEval otorgar mayor importancia al criterio de Calidad y utilidad del informe, por considerarlo de especial relevancia en la calificación de una evaluación. Por ello, se decidió otorgarle una puntuación de 20 puntos sobre el total. El resultado es una puntuación total máxima de 100 puntos para una evaluación.
- En definitiva, la ponderación de cada criterio y dimensión de la lista de metaevaluación quedó como sigue:

[5] Por ejemplo, puede que una evaluación de diseño no juzgue impactos o que una de impacto no juzgue la pertinencia.

Tabla 3. **Ponderación de criterios y dimensiones en la valoración cuantitativa**

CRITERIOS	SUBCRITERIOS/DIMENSIONES	PONDERACIÓN
1. Imparcialidad e independencia:		10 puntos
2. Equidad, participación y transparencia:		10 puntos
3. Credibilidad y exactitud:		50 puntos
3.1	Metas, objetos y modelo de cambio	10 puntos
3.2	Diseño y métodos	10 puntos
3.3	Procedimiento de selección	10 puntos
3.4	Análisis de datos	10 puntos
3.5	Criterios de evaluación	10 puntos
4. Calidad del informe y utilidad:		20 puntos
5. Exhaustividad y adecuación de la documentación		10 puntos
	PUNTUACIÓN TOTAL	100 PUNTOS

Fuente: Elaboración propia

- Finalmente, se establecieron estándares de cumplimiento de los criterios.
  - Cumplimiento alto: puntuación total mayor o igual que el 75% de la puntuación otorgada al criterio.
  - Cumplimiento intermedio: puntuación total mayor o igual que el 50% de la puntuación otorgada al criterio y menor del 75%.
  - Incumplimiento: puntuación total menor que el 50% de la puntuación otorgada al criterio.

  El sentido de establecer estos estándares de cumplimiento es, sobre todo, facilitar la sistemticidad de la respuesta a la pregunta central de esta metaevaluación sobre el indicador de FOCEVAL. Debe apuntarse que estos estándares de cumplimiento han sido fijados por el metaevaluador al no existir ningunos explícitos en FOCEVAL al ser la primera ocasión en que éstos se aplican.

- La aplicación a cada evaluación de esta lista y este método de calificación se recogió en Fichas de Registro

Tabla 4. Criterios y preguntas de metaevaluación

1.	IMPARCIALIDAD E INDEPENDENCIA
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados?

2.	EQUIDAD, PARTICIPACIÓN Y TRANSPARENCIA
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos?
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?
2.3	¿Los TdR fueron formulados de forma clara y comprensible?
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?

4.	CALIDAD DEL INFORME Y UTILIDAD
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?
4.5	¿Las conclusiones están basadas en los resultados?
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?

5.	EXHAUSTIVIDAD Y ADECUACIÓN DE LA DOCUMENTACIÓN
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos?

Fuente: FOCEVAL

3.	CREDIBILIDAD Y EXACTITUD
3.1	**Metas, objetos y modelo de cambio**
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?
	¿Se formularon hipótesis de impacto concluyentes?
	¿Se diferencia entre los diferentes niveles de impacto?
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?
3.2	**Diseño y métodos**
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) Entrevistas guiadas (número?); b) grupos focales (número?); c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.); d) otros?...
3.3	**Procedimiento de selección**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras?
3.4	**Análisis de datos**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?
3.4.2	¿Se cuantifican las respuestas siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?
3.5	**Criterios de evaluación**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa: a) es relevante/pertinente para el país, la región y el grupo meta b) logró las metas definidas c) es eficiente d) causó impactos (outcome e impacto) deseados y no deseados. e) causó impactos duraderos f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?

Limitaciones de la metaevaluación.

Este diseño de la metaevaluación presenta una serie de limitaciones que se centran en aspectos metodológicos que, si bien no inciden en su implementación, deben ser tenidos en cuenta para contextualizar sus hallazgos, conclusiones y recomendaciones, de forma que puedan ser interpretados de manera adecuada. Se destacan 4 limitaciones metodológicas:

1. **Los criterios, dimensiones y preguntas se han definido ex-post**. Esto significa que se juzga la calidad de las evaluaciones a partir de unos parámetros que no tuvieron disponibles los equipos de evaluación y de gestión de las evaluaciones durante su realización. Un ejemplo ilustrativo de su incidencia en la puntuación se aprecia con el criterio de Exhaustividad y adecuación de la documentación, que se traduce esencialmente en unos contenidos deseables de los anexos de los informes (además de su estructura y facilidad de lectura), un parámetro que de ser conocido de antemano por los equipos gestores de las evaluaciones y por los equipos evaluadores, debiera satisfacerse con solvencia, obteniendo así buenos registros en este criterio.

   La excepción en este sentido sería la evaluación de los servicios de prevención y atención del cáncer cuyo encargo se produjo una vez elaborada la primera versión de este informe (diciembre de 2016) y aunque dicha evaluación estaba en su fase final (entrega en febrero de 2017), se puede afirmar que se conocían los criterios y preguntas a partir de los que sería juzgada, situándose en una cierta posición de ventaja. A la luz de la Tabla 4, cabe interpretar que esto puede tener una incidencia menor en algún criterio como la exhaustividad de la documentación, algo en lo que parece razonable que el informe haya tenido en cuenta estos criterios.

2. **La metaevaluación se realiza exclusivamente a partir del análisis documental**. Cabe señalar que algunos de los estándares internacionales definidos en los criterios/preguntas se refieren al proceso evaluativo o a su contexto y pueden no estar reflejados en su justa medida en los informes de evaluación (producto). Entre las limitaciones que se derivan de centrar el trabajo en el análisis documental en exclusiva, destaca la dificultad que entraña valorar la calidad y utilidad del proceso evaluativo a partir de su Informe final, y sin contar con las percepciones cualitativas de los equipo de evaluación y de los equipos gestores de las mismas. En el caso de la valoración de la utilidad, esta aproximación no permite apreciar elementos relevantes del contexto evaluativo que pueden ser determinantes para el uso efectivo o real de las recomendaciones de las evaluaciones.

3. **La metaevaluación se realiza por un consultor individual y no por un equipo.** Esto tiene el riesgo de que no permite contrastar las valoraciones efectuadas, a través por ejemplo de un peer review, para corregir los sesgos que puede introducir del evaluador.

4. **La metaevaluación juzga la calidad y utilidad de Informes finales y TdR con un sesgo hacia los primeros.**

   - Por un lado, en la lista de criterios de metaevaluación las preguntas que refieren directamente -o indirectamente- a los TdR son limitidas, centrándose mayoritariamente en los informes finales.
   - Por otro lado, entre la documentación de base aportada para la metaevaluación, se dispuso de todos los informes finales de evaluación y de sus principales anexos (100%)[6], mientras que en el caso de los TdR se tuvo acceso a 3 documentos, por lo que el análisis corresponde a una muestra (30% del total[7]).

En cuanto a la implementación de la metaevaluación no se han presentado limitaciones, más allá de que la propia definición del encargo de los TdR que la regulan, establecía un plazo de realización muy ajustado en días de trabajo. El mismo día de la firma del contrato se dispuso de la documentación de base y de la lista de criterios y preguntas. Así, se dedicó un día de trabajo a la operacionalización de la lista de criterios, un día de trabajo para el análisis documental de cada evaluación y para su registro en fichas (10), un día para elaboración del informe y dos para su adaptación según la revisión del equipo FOCEVAL. La integración de la evaluación de los servicios de prevención y atención del cáncer supuso un día adicional de trabajo, más otro para actualizar el informe y confeccionar la presentación en diapositivas del informe.

[6] En el caso de Germinadora entre la documentación de base de la documentación de base aportada para realizar la metaevaluación no se encontraban los Estudios de Caso realizados, si bien formaron parte del producto de la evaluación y están disponibles en la web de MIDEPLAN.

[7] No se contempla aquí la evaluación de los servicios de prevención y atención del cáncer dado que no tenía TdR.

2. Metaevaluación de las evaluaciones

Este capítulo comienza con la exposición de una síntesis de las 11 evaluaciones evaluadas y se cierra con una apreciación de general de las evaluaciones a partir de una visión de conjunto y con una perspectiva comparativa que pretende identificar una caracterización así como identificar sus tendencias y patrones de comportamiento en torno a cada uno de los grandes criterios y subcriterios de evaluación aplicados. Finaliza con la emisión de las conclusiones que se derivan de los hallazgos de la metaevaluación.

2.1. Síntesis de las evaluaciones evaluadas

Los TdR estipulaban que el Informe debía tener un subcapítulo para cada evaluación evaluada, que incluyese:

1) Descripción del programa evaluado (un párrafo),
2) Descripción de la modalidad aplicada de implementación y apoyo a la evaluación (máximo media página)
3) Calificación cuantitativa de la evaluación y los respectivos términos de referencia según los criterios pre-establecidos.
4) Resumen de fortalezas y debilidades de la evaluación, con énfasis en la calidad general de la evaluación y el proceso.

En respuesta a este requerimiento, a continuación se presenta una síntesis de las 11 evaluaciones objeto de (meta) evaluación. Esta síntesis se nutre de la Ficha de registro de cada evaluación (Ver Anexo 4.4).

2.1.1 Evaluación 01: DRAT

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO
Distrito de Riego Arenal Tempisque (DRAT) en la Región de Chorotega (1981-2011)
2. INSTITUCIÓN QUE IMPLEMENTA
SENARA Servicio Nacional de Aguas, Riego y Avenimiento (Pág. 17-18 y Pág. 22 nota al pie 3); Junta Coordinadora del Distrito de Riego Arenal Tempisque (Pág. 17-18); INDER Instituto de Desarrollo Rural (antiguo IDA Instituto de Desarrollo Agrario) (Pág. 50) ; ICE Instituto Costarricense de Electricidad(Pág. 50)
3. MONTO DEL PROYECTO/ PROGRAMA
Etapa I: US$ 20.6 millones. Etapa II: US$ 40.8 millones. Etapa III: US$ 3 millones. El monto total del proyecto asciende a US$ 64.4 millones.

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO
Se trata de un programa público tendente a promover el riego en las actividades de desarrollo agropecuario y rural del país. En concreto el proyecto del Distrito de Riego Arenal Tempisque (DRAT) fue implementado en 1978 con el objetivo de poner bajo desarrollo agrícola con riego tierras en la Región Chorotega y así fomentar el desarrollo agropecuario mediante la agricultura intensiva; buscando incrementar la producción de alimentos básicos. Además, se buscó mejorar la situación socioeconómica de la región por medio de la creación de nuevas fuentes de empleo y el aumento en el ingreso de los productores.

EVALUACIÓN

5. EVALUADORES
Programa de Investigación en Desarrollo, Economía y Ambiente. División de Investigación y Desarrollo, CATIE (Centro Agronómico Tropical de Investigación y Enseñanza). Coordinación de Juan Robalino. Equipo evaluador: Gloriana Lang, Chelsia Moraes, Catalina Sandoval, Luis Vargas y Laura Villalobos. [Coordinación y Equipo Nacional- - Universidad de Costa Rica]

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN	
No especificado en el Informe final de evaluación. Sólo se indica que contó con el apoyo financiero de GIZ, tanto del programa FOCEVAL como del proyecto Agua para Guanacaste	**Fecha informe**	Enero 2014

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL
Evaluación externa (consultor nacional) con asistencia técnica de GIZ-FOCEVAL y la conformación de un Comité Revisor (GIZ-MIDEPLAN-SENARA-SEPSA-) que dirige la evaluación Esta evaluación estratégica de Gobierno fue realizada en el marco del Programa de Fortalecimiento de Capacidades en Evaluación (FOCEVAL) que se ejecuta en conjunto por parte de la Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ) y el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) por encargo del Ministerio Alemán de Cooperación Económica y Desarrollo (BMZ). La evaluación fue dirigida por un Comité Revisor conformado por miembros de GIZ (Jana Rauch), MIDEPLAN (Mario Gonzalez, Mainor Madrigal), SENARA (Juan Carlos Mora) y SEPSA (Ghiselle Rodríguez y Miriam Valverde). El representante de FOCEVAL estuvo presente durante la programación previa y en la fase del diseño metodológico de la evaluación, donde se definieron aspectos estratégicos como la cadena de resultados. El análisis documental relacionado con la variable “presión sobre áreas protegidas” estuvo a cargo de Jana Rauch (GIZ). Por tanto, de forma concreta, la participación de GIZ-FOCEVAL fue en calidad de asistencia técnica, haciendo aportes a nivel de recursos humanos y apoyo logístico.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA [Global y por Criterios]:	
86,9	**CALIFICACIÓN GLOBAL [Sobre 100]**
10	**Imparcialidad e independencia** [Sobre 10]
Se trata de una evaluación desarrollada por un equipo técnico externo a la organización que implementa el programa. Asimismo, las fuentes de información seleccionadas garantizaban la inclusión de pluralidad de perspectivas en el proceso evaluativo. Atendiendo a esto, se puede convenir que la evaluación garantizaba su imparcialidad e independencia.	
5,5	**Equidad, participación y transparencia** [Sobre 10]
A nivel de documentación hay que indicar que si bien el Informe final de evaluación es accesible al público, no se dispone de los TdR. En relación a la participación, habría sido necesario implicar a otros grupos de interés en el proceso de planificación, y a los stakeholders de forma previa al lanzamiento de la evaluación.	
46,1	**Credibilidad y exactitud** [Sobre 50]
El diseño metodológico de esta evaluación permite atribuirle un importante grado de exactitud y credibilidad. Tanto la planificación previa estableciendo una cadena de resultados, como la definición de las preguntas de evaluación y las hipótesis de procesos e impactos, permiten garantizar el desarrollo de un trabajo evaluativo de calidad. Tan sólo se aprecia como carencia el uso de un único método de análisis cualitativo (entrevistas en profundidad). Si bien se trata de una herramienta que aporta una información de calidad, habría resultado muy interesante complementarla con algún tipo de análisis donde se recabara la información con una perspectiva grupal (Grupos focales o estudios Delphi)	
18,8	**Calidad del informe y utilidad** [Sobre 20]
El análisis de los resultados efectuado por el equipo evaluador, así como la redacción de conclusiones y recomendaciones hacen que el Informe final sea un documento de evaluación de bastante calidad. Al mismo tiempo, hay que indicar que tanto la redacción del informe como su estructura lo dotan de un elevado nivel de utilidad. La no disposición de los TdR es lo que impide alcanzar una nota de 10 en este apartado.	
6,5	**Exhaustividad y adecuación de la documentación** [Sobre 10]
Si bien de forma global la documentación disponible es bastante exhaustiva y adecuada para los requerimientos que debe tener un informe final de evaluación, se echa en falta el documento de los TdR, así como información relativa al cronograma de trabajo de fase de campo y el presupuesto de la evaluación.	

10. RESUMEN DE LA FORTALEZAS Y DEBILIDADES DE LA EVALUACIÓN.

FORTALEZAS

- El equipo evaluador ha garantizado la independencia y pluralidad de perspectivas en el desarrollo del trabajo de evaluación.
- Se ha efectuado un trabajo de triangulación excelente, obteniendo información relevante tanto del ámbito cualitativo como cuantitativo. Al mismo tiempo, la triangulación ha permitido cotejar información obtenida a partir de distintas fuentes, pero relativa a un mismo aspecto analizado.
- El uso de técnicas de análisis cuantitativo que contemplan el diseño y aplicación de un contra-factual, supone contar con la posibilidad de poder inferir los resultados e impactos obtenidos a la existencia del programa de forma exclusiva.
- El enfoque con el que se ha realizado el análisis de los resultados obtenidos durante el desarrollo del trabajo de campo no solo ha estado vinculado a obtener una valoración de los impactos que están derivándose del programa evaluado, sino también a la mejora en su implementación y a la mejora de futuras evaluaciones.
- Tanto la estructura del informe como su redacción dotan al documento de un elevado nivel de calidad y utilidad.

DEBILIDADES

- Cuando se evalúa un programa público con el alcance y la importancia que tiene el DRAT para un sector estratégico como es la agricultura, se hace necesario contar con una mayor diversidad de agentes clave en la planificación y diseño de la evaluación. De forma concreta la planificación debería incluir agentes involucrados en el diseño y gestión del programa, así como beneficiarios del mismo.
- Aparece como carencia el no desarrollar algún tipo de actividad previa (jornada de sensibilización, presentación de la evaluación,…) a modo de instrumento que facilitara la participación de los stakeholders.
- El uso de técnicas de análisis cuantitativo como el DID, el PSM y el control sintético, es complejo y no ha conseguido garantizar la existencia de un vínculo directo y exclusivo entre los resultados e impactos registrados y la existencia del programa. *Quizás ante programas de la amplitud del DRAT puede ser más efectivo desarrollar el estudio de impacto a partir de modelos de evaluación que analizan la contribución del programa, frente a los de atribución utilizados en este trabajo. Sobre todo, si se tiene en cuenta que su carácter eminentemente técnico requiere capacidades instaladas para que sus resultados puedan trasladarse a la toma de decisiones*.
- El número de técnicas de recopilación de información cualitativa es limitado, sólo se utiliza la entrevista en profundidad.
- Existe una carencia relativa a información básica utilizada en el diseño metodológico. Si bien se hace referencia a un Estudio preliminar evaluativo (Imbach et al. 2013), se trata de una información de carácter relevante que no se incluye en el Informe final de evaluación.

2.1.2 Evaluación 02: EDU-INCLUSIVA

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO

Proyecto: Apoyo a la gestión pedagógica a centros educativos de calidad con orientación inclusiva

2. INSTITUCIÓN QUE IMPLEMENTA

Dirección de Desarrollo Curricular del Ministerio de Educación Pública (MEP) de Costa Rica, desde la acción de la Comisión de Apoyo a la Educación Inclusiva (CAEI), integrada por miembros de todos los departamentos de dicha Dirección y del Centro Nacional de Recursos para la Educación Inclusiva (CENAREC)

3. MONTO DEL PROYECTO/ PROGRAMA

No especificado en Informe final de evaluación (Anexo 9 únicos datos financieros). Según POA de la Dirección de Desarrollo Curricular para el 2011, se contó con un presupuesto de ¢38.159.200 millones para diversas actividades del proyecto (Pág.37).

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO

Este proyecto aborda la problemática de la exclusión del sistema educativo basándose en que se evidencia en un contexto caracterizado por la desarticulación entre el centro educativo y las necesidades del estudiantado, de las familias y de la comunidad, que genera educación desigual y excluyente para el estudiantado y lleva a peores registros de reprobación y a la deserción escolar, a tener menos oportunidades y a una menor calidad de vida. Para combatirlo propone una experiencia piloto en una muestra de centros educativos entre junio de 2009 y diciembre de 2012 (6 inician en 2009 y 4 se incorporan en 2011). Se trata de generar una transformación en estos centros que produzca prácticas y políticas institucionales para impulsar una cultura institucional inclusiva, mediante la integración de la comunidad educativa (familias, estudiantado, personal docente, técnico y administrativo y otros) y la aplicación del denominado "Índice de Inclusión". Este índice, constituye una estrategia metodológica conformada por un conjunto de materiales diseñados para apoyar a las escuelas en el proceso de cambio hacia comunidades escolares colaborativas que promuevan en todo el alumnado altos niveles de logro; y se desarrolla en 7 etapas: (1) Construcción y aprobación del proyecto; (2) Concienciación; (3) Capacitación a equipos de trabajo; (4) Estudio del contexto escolar; (5) Seguimiento a la propuesta; (6) Acompañamiento pedagógico a las instituciones educativas; y (7) evaluación y sistematización de la experiencia. El desarrollo de estas etapas determina la aplicación del Índice de inclusión en los centros educativos (objetivo general), de la que se espera que se deriven prácticas y políticas inclusivas, generando una cultura inclusiva (objetivos específicos).

EVALUACIÓN

5. EVALUADORES

CONSULTORA: Susan Francis; MIDEPLAN: Marjorie Valerín, María José Vega; MEP (Ministerio de Educación Pública): Mario Segura, Rafael Martínez; GIZ (Agencia Alemana de Cooperación Internacional): Mayela Zúñiga

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN	
No especificado en el Informe final de evaluación	**Fecha Informe**	Enero de 2014

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL

Evaluación piloto desarrollada con el apoyo de FOCEVAL I bajo la modalidad de Aprender-haciendo, mediante una consultora nacional y la conformación de un grupo de trabajo MIDEPLAN-GIZ-MEP. Este proyecto fue presentado por la CAEI como cierre de la intervención al concurso de Evaluaciones piloto que en noviembre – diciembre 2012 convocaron MIDEPLAN junto con la GIZ, en el marco del programa de FOCEVAL. Esto se plasmó en la conformación para tales efectos de un grupo de trabajo con la participación de representantes de MIDEPLAN quienes fungen como funcionarios de la Unidad de Acciones Estratégicas (UAE) del Área Evaluación y Seguimiento (AES), así como representantes de GIZ y del Ministerio de Educación Pública (MEP). Para su realización se contrató una consultora nacional experta en evaluación y educación. En el Informe hay referencias explícitas a esta modalidad de apoyo, y se indican las expectativas de participación de MIDEPLAN, FOCEVAL y GIZ vinculadas al desarrollo de la capacidad de evaluación (Anexo 2).

Entre los aspectos positivos, el equipo evaluador destaca:

> "… se considera valioso todo el aprendizaje logrado por el proceso evaluativo para todas las personas que participaron en él, la consideración del proceso de aprender haciendo establece oportunidades de reconstrucción de conocimientos en prácticas reales".

Entre los aspectos negativos o condicionantes, el equipo evaluador destaca:

> "…se requieren bases previas para asegurar que no se convierta en un proceso extenso en el tiempo y, por tanto, de alta inversión…" (Pág. 35)

Señala "la estrategia de articulación institucional para desarrollar la evaluación" como una de las dos principales limitaciones de la evaluación (junto a la identificación de la teoría de la intervención). Por ello, indica que "requiere la previsión de las capacidades de aprendizaje y trabajo en equipo interinstitucional". Concretamente se apunta que:

> "La modalidad de «Aprender haciendo» es un ejercicio muy complejo y costoso, lo cual implica altos niveles de compromiso, clarificación en los papeles y esfuerzos de coordinación importantes entre las instituciones participantes, para el desarrollo efectivo de evaluaciones sobre la gestión de las instituciones públicas" (Pág. 17 y 68).

Basta leer las lecciones aprendidas para comprender las tensiones que ha supuesto su aplicación en esta evaluación. El hecho de que no basta con un perfil experto en evaluación (se necesita también amplio conocimiento del funcionamiento del sector público y la política evaluada –Pág.17-) o de que la participación de representantes de la institución responsable del proyecto evaluado en el equipo evaluador genera tensiones que deben quedar resueltas:

> "En la planificación, la organización y la ejecución de procesos de evaluaciones públicas, la posibilidad de participación de miembros de la unidad ejecutora de la intervención que genera una serie de condiciones que deberán ser analizadas a priori: función que cumplen en el equipo evaluador, actividades que técnicamente están habilitadas para realizar, compromisos y productos de su participación en la evaluación. Además, es fundamental establecer en los procesos de negociación interinstitucional, los elementos de carácter ético y conceptual que implica integrar personal que participó de la ejecución de la intervención". (Pág. 16).

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA [Global y por Criterios]:

69,1	**CALIFICACIÓN GLOBAL [Sobre 100]**
7,5	**Imparcialidad e independencia**

Se califica como evaluación mixta pues aunque en el informe se califica como externa, se incluye el MEP en el equipo de evaluación y también se indica en el informe que el hecho de que la institución evaluada participara en la realización de la evaluación pudo ser un factor limitante, pero se contrarresta con una pluralidad de técnicas y fuentes, que ayuda a tener una visión global y completa.

10	**Equidad, participación y transparencia [Sobre 10]**

La evaluación destaca por la implicación y participación de los stakeholders, involucrados no sólo desde el rol convencional de suministradores de información (informantes clave), destacando en este sentido la participación activa de la CAEI en el diseño e implementación. También en transparencia los resultados son satisfactorios, ya que el informe está accesible en Internet, las preguntas de evaluación son claras y permiten hacerse una idea adecuada de los focos de la evaluación, aunque haya requerido procesar las inicialmente formuladas según se indica. El ítem que no aplicó fue la disposición de los TdR, no disponibles.

27,3	**Credibilidad y exactitud [Sobre 50]**

Las puntuaciones en las que se desglosa este criterio resultan desiguales en esta evaluación. Hay aspectos consolidados, como el análisis de los stakeholders y su uso en la evaluación, la correspondencia del diseño mediante estudios de caso a los intereses de la evaluación, o la pluralidad de métodos empleados que cubren lo cuantitativo y lo cualitativo, y el tratamiento aceptable y concluyente de los principales criterios que le son de aplicación. Así, sus mejores registros en credibilidad y exactitud se centran en el diseño y los métodos aplicados, y en su tratamiento y respuesta a los criterios de evaluación contemplados. Por el contrario, las dimensiones que reciben valoración intermedia están asociadas con la definición de metas, objetos y modelo de cambio, la especificación de los procedimientos de selección –a los que presta escasa atención-, y especialmente el análisis de datos, en el que le penaliza que no contase con datos del monitoreo del programa –limitados según se indica- ni recurriese a fuentes de datos externas para facilitar la contextualización y comparación.

18,8	**Calidad del informe y utilidad [Sobre 20]**

En este criterio la evaluación obtiene sus mejores registros. Presenta la más alta valoración en la mayoría de ítems, siendo su principal debilidad la escasa atención que presta a las cuestiones de fiabilidad en los hallazgos o conclusiones presentados.

5,5	**Exhaustividad y adecuación de la documentación [Sobre 10]**

La exhaustividad y adecuación de la documentación destaca en cuanto que presenta una estructura adecuada y lectura apoyada en tablas y gráficos así como en la inclusión en anexos de los instrumentos aplicados para recabar información. No obstante, sus limitaciones se centran en no presentar anexos los TdR y en la mejorable especificación en anexos del trabajo de campo realizado.

10. RESUMEN DE FORTALEZAS Y DEBILIDADES DE LA EVALUACIÓN

FORTALEZAS

- Realiza un amplio trabajo de campo cualitativo, con alta variedad de stakeholders y participantes como informantes clave.
- La participación como eje relevante del proceso, más allá de la consulta de agentes clave, especialmente desarrollada con CAEI.
- Destaca la construcción participativa de preguntas y la delimitación del encargo evaluativo con la CAEI.
- El establecimiento de recomendaciones por actor supone una importante orientación práctica que debe facilitar su aplicación así como el establecimiento de un plan de acción para el seguimiento de la implementación de estas recomendaciones.
- Sistematicidad: la operacionalización de los criterios evaluativos; el establecimiento de categorías de análisis y parámetros para su valoración, definiendo elementos, patrones y estructuras para cada una de ellas, y su operacionalización en indicadores (Anexo 5) refuerza el diseño de la evaluación, y el tratamiento sistemático de la abundante información cualitativa recabada.
- Esfuerzo por introducir entre los hallazgos frases literales entrecomilladas de las personas entrevistas, que contextualizan acertadamente aquello que se expone (entrevista con la CAEI, entrevistas grupales y algún estudio de caso).

DEBILIDADES

- El diseño de la evaluación apunta enfoques e indicadores para los que se aprecia un limitado tratamiento en la evaluación.
  - El modelo CIPP que anuncia como anclaje (contexto, insumos, procesos, productos), considerándose acertado, no se aprecia con claridad su uso más allá de la exposición de una "cadena de resultados" excesivamente genérica.
  - Los indicadores del Índice de inclusión se presentan en anexo pero no se utilizan para estudiar y comparar los casos.
  - Centra el uso de métodos y técnicas en estudios de caso (acertado teniendo en cuenta que la sistematización era un objetivo), si bien no se incorporan ni anexan, cuando se nutre de ellos, y las referencias a sus resultados son difusas o dificultan la trazabilidad de los hallazgos, conclusiones y recomendaciones.
  - Se otorga importancia al estudio en profundidad de tres centros educativos (Pág.30) pero no se especifica en qué consiste ni se aprecia de forma diferencia qué resultados produjo para la evaluación.
  - La definición operacional de los criterios, parámetros e indicadores (Anexo) con escalas de valoración no se aprecia de manera clara y sistemática en la exposición de hallazgos. Es probable que hayan tenido mayor uso y reflejo en los estudios de caso.
- Entre las limitaciones de la evaluación no se referencia la disponibilidad de información o la precisión del alcance que sí aparecen como Lecciones aprendidas, en el sentido de que no contribuían a alcanzar los propósitos de la evaluación. Se intuye por el contrario, aunque no se explicita, que la falta de información disponible, y el escaso soporte documental de los proyectos, han sido limitaciones importantes que se han enfrentado, a no ser que se abordasen en los estudios de caso.
- Aprovechamiento desigual de la información recabada, especialmente la referida a los diversos cuestionarios.
- El enjuiciamiento a partir de los criterios (eficacia, pertinencia, sostenibilidad, calidad.) queda diluido en otras lecturas del proyecto (especialmente en los dos niveles de análisis y los ámbitos de indagación establecidos), encontrando diferentes acepciones de los criterios (una las que define explícitamente y otra la que aplica, especialmente en pertinencia). Que los ámbitos de indagación establecidos en el diseño vertebrasen la exposición de recomendaciones hubiese facilitado el examen de su trazabilidad con los hallazgos, pero ha restado utilidad al importante tratamiento sistemático de información cualitativa.
- La escala de valoración de resultados del cuestionario a alumnos esta balanceada: escala Likert de tres categorías, siendo la central "Parcialmente de acuerdo" (para estudiantes de ciclo diversificado) y "De acuerdo" (para estudiantes de I y II Ciclo)

2.1.3 Evaluación 03: GERMINADORA

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO

PROYECTO INTERINSTITUCIONAL GERMINADORA DE DESARROLLO ORGANIZACIONAL EMPRESARIAL, ASOCIATIVO Y COMUNITARIO.

2. INSTITUCIÓN QUE IMPLEMENTA

IMAS e Instituto Nacional de Aprendizaje (INA) asumen la elaboración de una versión interinstitucional del proyecto, siendo IMAS la Unidad coordinadora. Su ejecución es interinstitucional, están el Comité Técnico Interinstitucional que se encuentra formado por enlaces de las instituciones implicadas en el proyecto: Universidad Nacional de Costa Rica (**UNA**), Instituto Mixto de Ayuda Social (**IMAS**), Instituto Nacional de Aprendizaje (**INA**), Banco Popular y de Desarrollo Comunal (**BPDC**), Junta de Desarrollo Regional de la Zona Sur (**JUDESUR**) y el Movimiento Cooperativo: Instituto de Fomento Cooperativo (**INFOCOOP**), Centro Nacional de Educación Cooperativa (**CENECOOP R.L**), Consejo Nacional de Cooperativas (**CONACOOP**).

3. MONTO DEL PROYECTO/ PROGRAMA

No se especifica. Se indica que no existe un presupuesto para este proyecto y no se puede determinar la inversión total realizada.

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO

Antecedentes. El proyecto evaluado, desarrollado en la Región de Brunca de Costa Rica, es fruto de diversas iniciativas iniciadas en 2005 por la Escuela Planificación y Promoción Social (EPPS), de la Universidad Nacional (UNA), para impulsar procesos de desarrollo local y diversificación de la actividad productiva de la región. En 2010 se desarrolló un "Laboratorio Organizacional de Curso (LOC)" para formar Técnicos en Desarrollo Cooperativo Comunitario (TDCC). La idea era que los graduados consolidaran una empresa de TDCC, que en una fase posterior, pudiesen ofrecer sus servicios como directores de LOT, y formar Auxiliares en proyectos de inversión (API), en la región Brunca. En 2011, IMAS y el Instituto Nacional de Aprendizaje (INA), asumen la elaboración de una versión interinstitucional del proyecto, reuniendo estos antecedentes y con el fin de impulsar una visión compartida que permitiera gestionar recursos, y por esto es formulado con el nombre; "Proyecto Interinstitucional Germinadora de Desarrollo Organizacional, Empresarial, Asociativo y Comunitario", que es el que se encuentra en proceso de evaluación.

El "Proyecto Germinadora", en su versión interinstitucional, está orientado al desarrollo empresarial y organizativo que repercuta en la dinamización de la economía local y en el fortalecimiento de la acción ciudadana para la promoción del desarrollo endógeno de los espacios locales y de sus poblaciones más vulnerables. Es una iniciativa que combina la capacitación, promoción y seguimiento, para generar ingresos a través de la iniciativa empresarial identificada masivamente en zonas en condiciones de pobreza. Encuentra su fundamento en la denominada "Metodología de Capacitación Masiva (MCM)" que procura la estimulación de la conciencia organizativa de la población participante mediante el desarrollo de las capacidades de organización empresarial y ciudadana articuladas en los denominados "Laboratorios Organizacionales –LOT-". Se dirige a población adulta, especialmente de familias cuyos jefes y jefas de hogar perdieron la oportunidad de la educación formal, y se ven desplazados por las altas exigencias del mercado y los avances tecnológicos. Por ello, busca una alfabetización empresarial y técnica que permita a las personas emprendedoras, ubicar y desarrollar proyectos en nichos de mercado que estimulen los encadenamientos locales, enriquezca la actividad económica local, genere ingresos básicos a las familias y, al mismo tiempo, desarrollen capacidades empresariales.

EVALUACIÓN

5. EVALUADORES

Equipo de Evaluación: Kemly Camacho (Coordinadora, evaluadora principal); Rosaura Herrera, Ericka Valerio (MIDEPLAN); Carol Peña (GIZ). Se contó con el apoyo de Adriana Sánchez (Enlace institucional IMAS)

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN	
No se especifica en el informe y no se dispone de TdR.	**Fecha Informe**:	Febrero de 2014

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL

Evaluación piloto apoyada por FOCEVAL I bajo la modalidad Aprender-haciendo y ejecutada por consultora nacional apoyada por MIDEPLAN-GIZ-FOCEVAL con enlace institucional del IMAS (responsable del programa evaluado).

El informe indica que "Este proceso ha sido particular ya que está acompañado de un objetivo de fortalecimiento de capacidades en evaluación para las personas que han integrado el equipo de trabajo y que son representantes del IMAS y del Programa FOCEVAL (MIDEPLAN y la GIZ), bajo la modalidad "aprender haciendo"". (Pág.6 y 39)

"Dada la naturaleza de esta evaluación … se realiza por medio de un evaluador externo, dos funcionarias de la Unidad de Acciones Estratégicas de MIDEPLAN, una funcionaria de la GIZ-FOCEVAL y una contraparte técnica institucional del IMAS, con el objetivo de aplicar la metodología de aprender haciendo de modo que los funcionarios desarrollen capacidades en evaluación. (Pág.39)

Aspectos positivos:

- "Se realizó un muy buen proceso aprender-haciendo sobre todo en la parte del diseño evaluativo". (Pág. 134)
- "Al aplicar una metodología de aprender-haciendo se fomenta el desarrollo de capacidades en los representantes institucionales de MIDEPLAN, e IMAS de tal forma que puedan replicar dicha experiencia en futuras evaluaciones de proyectos estratégicos, y en la mejora de su desempeño en la ejecución de sus competencias institucionales". (Pág. 45)

Limitaciones o condicionantes:

- "El proceso de aprender-haciendo provoca que el proceso evaluativo sea más lento y se deba someter el avance… de acuerdo con las capacidades formadas y el nivel de asimilación de contenidos que demuestre parte del equipo evaluador." (Pág. 47)
- "El equipo evaluador no está únicamente dedicado a la evaluación del Proyecto germinadora… podría limitar su disponibilidad y flexibilidad para asumir más allá de medio tiempo laboral previsto para invertir en la evaluación. (Pág.47)
- "Debido a la extensión de esta evaluación se fue complicando la disponibilidad de tiempo de todo el equipo para que este continuara involucrado el medio tiempo de cada integrante como se había propuesto originalmente. Esto produjo que el proceso se alargara más de lo esperado". (Pág. 134)

Reflexiones finales sobre este proceso evaluativo. (Pág. 134)

- La reconstrucción de la teoría del programa del proyecto se percibe como una labor muy importante que consume muchos recursos y tiempo y que fue posible "por el aporte de FOCEVAL y al contexto de aprendizaje y de pilotaje en el que se enmarca esta experiencia. Sin embargo, los recursos dedicados a este proceso evaluativo fueron insuficientes, se ha requerido de mucho tiempo extraordinario aportado por la evaluadora principal y la GIZ."
- En términos de las capacidades de evaluación, además de la necesidad de formular mejores proyectos y fortalecer el seguimiento de los mismos es indispensable para los equipos de evaluación fortalecer las capacidades de investigación y garantizar el manejo de las herramientas cuantitativas y cualitativas y su combinación. Sin una fortaleza investigativa se reduce las posibilidades de una investigación rigurosa que alimente una buena toma de decisiones y el mejoramiento de las acciones.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA: [Global y por Criterios]:

78,0	**Calificación global** [Sobre 100]
10	**Imparcialidad e independencia** [Sobre 10]

La evaluación presenta la máxima puntuación en los dos ítems que desagregan este criterio, pues es de tipo externo (IMAS, con responsabilidad en el proyecto, funge como enlace institucional con el equipo evaluador) y utilizó fuentes que permiten capturar una pluralidad de perspectivas en el levantamiento y análisis de datos así como en la presentación de resultados.

7,5	**Equidad, participación y transparencia** [Sobre 10]

La evaluación obtiene una buena calificación en este criterio pues participan activamente en su planificación e implementación grupos interesados (especialmente la Comisión Interinstitucional) a través de talleres para la reconstrucción de la teoría de la intervención y se conforman círculos hermenéuticos para valorar los hallazgos con stakeholders. Aunque no se especifica que se llevasen a cabo talleres de inicio con grupos meta y stakeholders, se contemplaron los talleres citados para el diseño de la teoría de la intervención. Las preguntas de evaluación son claras y adaptadas al contexto. No se dispone de TdR por lo que no se puede juzgar su claridad y comprensión. Y el informe está accesible en internet, no así el resumen ejecutivo.

37,5	**Credibilidad y exactitud** [Sobre 50]

Destaca la evaluación en la delimitación y descripción clara de sus metas, objeto y modelo de cambio (8,7), presentando un completo análisis de los grupos de interés, su elaboración de la teoría de la intervención, diferenciando niveles de impacto, y presenta una descripción exhaustiva del contexto evaluativo.

Presenta una buena valoración en términos de credibilidad y exactitud en el diseño y los métodos aplicados (8,0), pues su diseño se corresponde con los intereses de la evaluación, se aplicaron diferentes métodos que cubren aspectos cuantitativos y cualitativos. En cuanto a la consideración de lo contrafactual, lo contempló aunque descartó un diseño cuasiexperimental por problemas con la información y construyó en su defecto una comparación de poblaciones API y no API mediante encuestación.

En cuanto a procedimientos de selección presenta una valoración intermedia (5), pues sólo explicita en algunos casos los criterios aplicados (encuesta y sondeo) pero no en el resto de técnicas cualitativas; y en la muestra de la encuestas siendo representativa se contempla un error muy elevado (7%) y establece un sondeo a conveniencia con un número limitado de No API.

En cuanto al análisis de datos obtiene una calificación suficiente (6,25), pues triangula la información que se recaba a través de diferentes técnicas, se cuantifican a veces las respuestas que se exponen, hace un uso limitado de la información de monitoreo y recurre en algunos casos a fuentes externas para la comparación. Finalmente, presenta credibilidad y exactitud en el tratamiento y enjuiciamiento de los criterios de evaluación (10) que le son de aplicación (a excepción de los relativos a impactos y sostenibilidad).

16,5	**Calidad del informe y utilidad** [Sobre 20]

La evaluación de la calidad y utilidad presenta resultados notables (8,71), pues el informe da respuesta adecuada a todas las preguntas de evaluación, describe de forma balanceada sus fortalezas y debilidades, diferencias valoraciones descriptivas y valorativas y existe trazabilidad entre hallazgos, conclusiones y recomendaciones. Las recomendaciones que emite son relevantes y su grado de concreción es mejorable así como su orientación a la acción (priorización, identificación de destinatarios). Como debilidad se destaca la ausencia de un resumen ejecutivo en el propio informe, lo que merma notablemente sus registros en utilidad y facilidad de uso.

6,5	**Exhaustividad y adecuación de la documentación** [Sobre 10]

En el criterio de exhaustividad y adecuación de la documentación la evaluación recibe una calificación intermedia (5,5), en tanto que no presenta en anexos los TdR y en estos sólo incluye algunos de los documentos y datos que especifican el trabajo de campo realizado, aunque si anexa los instrumentos que aplica.

Presenta una estructura adecuada y su lectura está facilitada por gráficos y tablas aunque no con resúmenes de los capítulos

10. RESUMEN DE FORTALEZAS Y DEBILIDADES DE LA EVALUACIÓN

FORTALEZAS

- Amplio y sistemático trabajo de campo cualitativo, abarcando un amplio espectro de stakeholders.
- Sistematicidad del proceso de enjuiciamiento y análisis evaluativo, a partir del establecimiento de categorías de análisis y variables/indicadores en el diseño, que durante la implementación ayudan a vertebrar la exposición de hallazgos, facilitando el examen de la trazabilidad de hallazgos, conclusiones y recomendaciones por pregunta evaluativa.
- La sistematicidad del tratamiento de la información cualitativa recabada.

DEBILIDADES

- El amplio margen de error en la muestra de API (+-7%) junto a la realización de un sondeo con una muestra a conveniencia, hacen que algunas de las conclusiones que se emiten con base en la comparación de ambas, deban tomarse con cautela
- La conformación de la comparación API con No API a partir de sondeo a conveniencia deriva en una muestra de 41 personas que, aunque se explicitan sus fines, requiere cautela al establecer comparaciones.
- Al ser uno de los intereses de la evaluación conocer la replicabilidad del proyecto en otras regiones y cantones, el diseño evaluativo no presenta la validez externa necesaria, si bien justifica la imposibilidad de haberlo contemplado.

- En la exposición del objeto evaluativo se introducen conceptos que hubiesen requerido una explicación previa (Laboratorios organizaciones –LOT, ¿en qué se diferencia de los LOC?, curso preprofesionales, etc.-) que si bien el lector encuentra a lo largo del informe o en Anexos, hace más compleja la primera lectura del documento exigiendo un esfuerzo añadido.

2.1.4 Evaluación 04: PND

PROGRAMA / PROYECTO EVALUADO
1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO
Evaluación del "Diseño Metodológico y Proceso de Elaboración del PND 2011 – 2014", de MIDEPLAN.
2. INSTITUCIÓN QUE IMPLEMENTA
Ministerio de Planificación Nacional y Política Económica (MIDEPLAN): responsabilidad principal de la elaboración del PND. (Art. 4º de la Ley No 5525) ; Implanta normas de asesoría, información y coordinación necesarias con el resto del Sistema Nacional de Planificación (SNP)[8]
3. MONTO DEL PROYECTO/ PROGRAMA
PND:5,2 billones de colones (Pág. 20)[9]

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO
El objeto de esta evaluación es el Plan Nacional de Desarrollo (PND) "María Teresa Obregón Zamora" de Costa Rica 2011-2014. No obstante el interés evaluativo se centra en mejorar el diseño metodológico y el proceso de elaboración del Plan Nacional de Desarrollo (PND) 2015-2018 a partir de la experiencia del PND 2011-2014 y proponer una nueva metodología para orientar la planeación hacia resultados más estratégicos y evaluables, que permita una vinculación con marcos presupuestarios de corto-mediano plazo y que facilite una participación más amplia de diversos actores representativos sociales e institucionales. El PND 2011-2014 presenta una estructura que busca enlazar las políticas y metas nacionales con las políticas, metas y acciones estratégicas sectoriales. MIDEPLAN tiene la responsabilidad principal de la elaboración del PND. Para ello debe implantar normas de asesoría, información y coordinación que sean necesarias con el resto del Sistema Nacional de Planificación (SNP), el cual deberá prestarle toda la cooperación técnica requerida. Así, MIDEPLAN emite una metodología para la elaboración del PND, con el propósito de que los actores, definan metas, incorporen programas, proyectos, asignen recursos y definan responsables y así ordenar los procesos de las entidades, al tiempo que permita el seguimiento y evaluación, la transparencia y la rendición de cuentas. Los avances de monitoreo y seguimiento al PND 2011-2014 muestran brechas entre los objetivos planteados y el resultado final que se elabora a partir de la interacción de los actores que intervienen en el proceso de construcción del PND. Incluso, la Contraloría General de la República ha señalado en sus informes la necesidad de mejorar los indicadores para la medición de la eficiencia y que se presentan divergencias entre indicadores, metas y estimación de recursos al comparar lo establecido en el PND 2011-2014, recomendando "mejorar la vinculación entre la planificación y la presupuestación (…) estableciendo procedimientos para la vinculación en los momentos claves de los procesos, como es la elaboración del PND".

EVALUACIÓN		
5. EVALUADORES		
Instituto IDEA International		
6. MONTO PARA LA EVALUACIÓN	**7. FECHA DE REALIZACIÓN**	
No se especifica en el informe y no se dispone de TdR.	**Fecha Informe**:	1 Agosto 2013
8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL		
Evaluación externa (Consultoría Internacional), apoyada por FOCEVAL I[10] El hecho de que la evaluación se realizase bajo esta modalidad de evaluación externa hace que en el informe no se preste atención al desarrollo de capacidades que se haya podido derivar del proceso, ni tampoco otras que puedan inferirse como aprendizajes derivados del proceso, fortalezas o limitaciones. Simplemente se limita a exponer que FOCEVAL, junto a MIDEPLAN, patrocinan esta evaluación y que ambas organizaciones supervisaron el desarrollo del trabajo de campo de la evaluación y sus principales instrumentos aplicados. Tampoco hay referencias a si se conformó una estructura para la gestión de la evaluación con MIDEPLAN o FOCEVAL-GIZ.		

[8] Forman el SNP: a)MIDEPLAN; b) Las unidades u oficinas de planificación de los ministerios, instituciones descentralizadas y entidades públicas locales y regionales; y c) Los mecanismos de coordinación y asesoría, tales como consejos asesores, comités interinstitucionales, comisiones consultivas y otros

[9] En el informe se indica que esta estimación por ejes que contiene el PND no es consistente con la estimación presupuestal de los cuadros de Lineamientos de Políticas Sectoriales del Anexo 1

[10] Aunque los TdR de esta metaevaluación indicaban que esta evaluación fue apoyada bajo la modalidad de aprender-haciendo, se realizó bajo la modalidad de evaluación externa.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA: [Global y por Criterios]:

73,1	**Calificación global** [Sobre 100]
10	**Imparcialidad e independencia** [Sobre 10]

La evaluación obtiene la máxima puntuación en este criterio al ser de tipo externo en la que además se considera que las fuentes seleccionadas permiten capturar una pluralidad de perspectivas en sus distintas etapas.

6,5	**Equidad, participación y transparencia** [Sobre 10]

La evaluación obtiene una calificación suficiente en este criterio (pues participan activamente en implementación grupos interesados –si bien como informante clave- (aunque no se aprecia la participación en el diseño). En cuanto a las preguntas de evaluación son claras y adaptadas al contexto. No se dispone de TdR por lo que no se puede juzgar su claridad y comprensión. Y tanto el informe como el resumen ejecutivo que aquél contiene están accesibles en internet.

32,5	**Credibilidad y exactitud** [Sobre 50]

La evaluación obtiene una valoración notable en este criterio (7), pues está descrita la meta/objetivo de la evaluación, describe el objeto evaluado aunque sea en algunos aspectos (no en todos) y describe exhaustiva el contexto evaluativo. Le penaliza en este sentido que no se aprecia que realizase un taller inicial con stakeholders para facilitar el proceso evaluativo.

Presenta una valoración notable en términos de credibilidad y exactitud en el diseño y los métodos aplicados (8,75), pues su diseño se corresponde con los intereses de la evaluación y se aplicaron diferentes métodos que cubren aspectos cuantitativos y cualitativos. En cuanto a la consideración de lo contrafactual, esta evaluación no lo contempla, si bien se considera que no aplica teniendo en cuenta los intereses evaluativos.

En cuanto a los procedimientos y criterios de selección aplicados (5) no se explicitan en todos los casos y en el caso de la encuesta presenta insuficiente información para juzgar la bondad de la muestra.

En cuanto al análisis de datos obtiene una calificación intermedia (5), pues triangula la información que se recaba a través de diferentes técnicas y lo expone con claridad y hace un uso de la información de monitoreo. Le penaliza en este sentido que sólo a veces cuantifica las respuestas que se exponen (teniendo que recurrir a los anexos en el caso de la encuesta) y no recurre a fuentes externas para la comparación durante el análisis (sólo algo para contextualizar). Y, finalmente, presenta credibilidad y exactitud en el tratamiento y enjuiciamiento de los criterios de evaluación (9) que le son de aplicación (a excepción de los relativos a impactos y eficiencia).

18,8	**Calidad del informe y utilidad** [Sobre 20]

La evaluación de la calidad y utilidad presenta resultados notables (9,38), pues el informe da respuesta adecuada a todas las preguntas de evaluación, describe de forma balanceada sus fortalezas y debilidades, diferencia moderadamente valoraciones descriptivas y valorativas y existe trazabilidad entre hallazgos, conclusiones y recomendaciones. Las recomendaciones que emite son relevantes y su grado de concreción adecuado (reforzado por sus detallados plan de acción y estrategia de socialización), junto a su articulación en una propuesta de diseño para el nuevo PND. Además contempla un resumen ejecutivo como parte del informe.

5,3	**Exhaustividad y adecuación de la documentación** [Sobre 10]

En el criterio de exhaustividad y adecuación de la documentación la evaluación recibe una calificación suficiente (6,5), en tanto que no presenta en anexos los TdR y en estos sólo incluyen algunos de los documentos y datos que especifican el trabajo de campo realizado, aunque si anexa los instrumentos que aplica. Presenta una estructura adecuada y su lectura está facilitada por gráficos y tablas aunque no con resúmenes de los capítulos

10. RESUMEN DE FORTALEZAS Y DEBILIDADES DE LA EVALUACIÓN

FORTALEZAS

- Profundo y detallado examen técnico del diseño metodológico del PND en términos de racionalidad, pertinencia y coherencia interna y externa.
- Detallado Plan de acción y concreta y específica Estrategia de socialización como instrumentos facilitadores de la aplicación de las recomendaciones de esta evaluación.
- Importante trabajo de campo cualitativo con una buena sistematización de información que se expone con claridad para los lectores, lo cual debe maximizar su utilidad final, facilitando su aplicación.
- Exhaustivo trabajo de triangulación de la información, asentando en la introducción preguntas comunes a través de diversas técnicas, expuesto con claridad y de fácil interpretación.

DEBILIDADES

- Concede protagonismo en la triangulación al criterio experto del equipo evaluador (una de las cuatro fuentes trianguladas) sin exponer con claridad los hallazgos que lo justifican ni las prácticas internacionales que se refieren como base para su juicio (salvo excepciones). Si puede constituir una fuente relevante el dictamen experto que realizan sobre la calidad de la planificación si bien esto no se presenta así ni tampoco está sistematizado de manera que el lector pueda encontrar las fuentes y la trazabilidad de sus valoraciones. Esto resta rigor técnico al juicio final de la evaluación.
- Prima en exceso la claridad y facilidad de uso y lectura del informe, descuidando aspecto técnicos y metodológicos relevantes, como ir presentando los datos que sustentan los hallazgos de la evaluación, lo cual exige al lector a un tratamiento y revisión de los anexos para conocer las bases que sustentan el juicio de valor que se emite (esto ocurre especialmente en la encuesta), descuidando también el hecho de ir referenciando continuamente la fuente de la que se nutre los datos y hechos que se exponen.

2.1.5 Evaluación 05: REDCUDI

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO

Evaluación del diseño, gestión y ejecución del beneficio "Alternativas de Atención y Desarrollo de la Niñez", que atienden a los niños y a las niñas que reciben el subsidio del IMAS en el marco del Programa Red Nacional de Cuido y Desarrollo Infantil, período 2014 y 2015

2. INSTITUCIÓN QUE IMPLEMENTA

Instituto Mixto de Ayuda Social (IMAS) y Secretaría Técnica de la Red Nacional de Cuido y Desarrollo Infantil

3. MONTO DEL PROYECTO/ PROGRAMA

No se especifica. Las únicas referencias son: "En el 1er semestre 2015, el IMAS otorgó subsidios … para sufragar el costo de atención en Alternativas de Atención a la Niñez … para una inversión de ¢7.368.941,3 miles de colones"; y "la meta presupuestaria para 2015 ascendió a ₡21,721,039,200,00

4. Descripción del proyecto/programa evaluado

El beneficio del IMAS consiste en la entrega de un subsidio a las familias para que sus hijos asistan a las Alternativas de Atención y Desarrollo de la Niñez (AADN), prioritariamente niños y niñas de cero a siete años, que viven en condiciones de pobreza, pobreza extrema y con vulnerabilidad social; no obstante, según necesidades específicas y disponibilidad presupuestaria, se pueden incluir niños y niñas hasta de doce años de edad que se encuentran en pobreza extrema y pobreza.

Las transferencias económicas deben ir acompañadas por procesos o acciones de promoción y orientación individual o familiar, que promuevan la movilidad social ascendente.

Y las modalidades de AADN ofrecen diferentes servicios y horarios de atención a padres, madres o encargados legales de niños y niñas prioritariamente en condiciones de pobreza son: Centros de Cuido y Desarrollo Infantil (CECUDI) - Oferta de servicio prestada por alternativas de cuido municipales-; Hogares Comunitarios - Oferta de servicios que presta una Madre Comunitaria-; y Centros infantiles - Oferta de servicio prestada por alternativas de cuido privadas o mixtas, incluidas cooperativas, asociaciones de desarrollo, asociaciones solidaristas y pequeñas empresas-

EVALUACIÓN

5. EVALUADORES

Centro de Investigación y Capacitación en Administración Pública, Universidad de Costa Rica (CICAP-UCR)

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN	
¢50.000.000,00 (50 millones de colones) [Aprox. 85.600 €]. De este presupuesto, que correspondía a dos fases de la evaluación, únicamente se ejecutó y canceló lo correspondiente al compromiso de MIDEPLAN (35 millones de colones), pues la segunda fase (IMAS) no fue ejecutada.	**Fecha Informe**	Junio, 2016

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL

Evaluación externa (Consultoría nacional–CICAP-UCR) en el marco de la ANE con apoyo de FOCEVAL II mediante la modalidad de Padrinazgo (asesor internacional) y la conformación de una Equipo Gestor y un Equipo Tomador de Decisiones (MIDEPLAN-IMAS-.ST REDCUDI). Según apuntan los TdR (Pág.15 y 16):

> "Se contará con el apoyo de FOCEVAL a lo largo de todo el proceso de la evaluación, en donde se aportará la colaboración de un experto internacional en la figura de padrino el cual estaría contribuyendo a la calidad del proceso de evaluación, reforzando las capacidades de los actores implicados en dicho proceso. La modalidad del "padrinazgo de una evaluación" se basa en la figura de un consultor internacional experto en evaluación (padrino) que acompaña el proceso de evaluación asesorando técnica y metodológicamente al Equipo de Gestión y a través de éste, al Equipo Evaluador.
>
> Está previsto que el asesoramiento del padrino tenga lugar desde la etapa de diseño de la evaluación, de manera que pueda acompañar en la toma de decisiones clave para el resto del proceso: delimitación del objeto de la evaluación, definición las preguntas claves, diseño metodológico, elaboración de términos de referencia para contratar al equipo evaluador, etc. El apoyo del padrino se extenderá hasta concluida la etapa de diseminación de los resultados y la elaboración de una estrategia para la aplicación de las recomendaciones de la evaluación ("management response")".

Al margen de esta delimitación de la modalidad que hacen los TdR, en el informe no se aprecia ninguna referencia expresa a las bondades o limitaciones que ha supuesto esta experiencia, ni directa ni indirectamente. Tan sólo cabe señalar que se apunta que la evaluación con una estructura de gestión (Equipo gestor y Equipo tomador de decisiones).

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA: [Global y por Criterios]:

67,6	**CALIFICACIÓN GLOBAL** [Sobre 100]
7,5	**Imparcialidad e independencia**

Arroja buena puntuación en Imparcialidad e independencia (7,5) pues se trata de una evaluación externa en la que además se considera que las fuentes seleccionadas permiten capturar una pluralidad de perspectivas en sus distintas etapas.

6	**Equidad, participación y transparencia** [Sobre 10]

La evaluación obtiene una calificación suficiente en este criterio (6), pues participan activamente en el diseño grupos interesados –mediante el taller de diseño de la teoría de la intervención- (aunque no se aprecia la participación activa en la implementación más que como informantes). En cuanto a las preguntas de evaluación si bien son claras se consideran complejas (cada una encierra más de un criterio evaluativo), y los TdR presentan una estructura adecuada y una extensión considerable que hubiese requerido mayor clarificación del objeto o un índice entre otros aspectos. Y tanto el informe como el resumen ejecutivo que aquél contiene están accesibles en internet.

30	**Credibilidad y exactitud** [Sobre 50]

La evaluación obtiene una valoración suficiente en este criterio (6,2), pues está descrita la meta/objetivo de la evaluación, describe con precisión el objeto evaluado y describe relativamente el contexto evaluativo, presenta un completo análisis de los stakeholders y una detallada teoría de la intervención.

Presenta una valoración notable en términos de credibilidad y exactitud en el diseño y los métodos aplicados (7,2), pues su diseño se corresponde con los intereses de la evaluación (con la salvedad del abordaje de la ejecución del Beneficio) y se aplicaron diferentes métodos que cubren aspectos cuantitativos y cualitativos. En cuanto a la consideración de lo contrafactual, esta evaluación no lo contempla, si bien se considera que no aplica teniendo en cuenta los intereses evaluativos.

En cuanto a los procedimientos y criterios de selección aplicados (5) no se explicitan en todos los casos y en el caso de la encuesta presenta una muestra adecuada con elevados errores en algunos casos.

En cuanto al análisis de datos obtiene una calificación intermedia (5,25), pues triangula la información que se recaba a través de diferentes técnicas (aunque no generalizadamente) y cuantifica las respuestas obtenidas cuando es posible, incluso a veces en las técnicas cualitativas. Le penaliza en este sentido el casi inexistente uso de datos del monitoreo y de fuentes externas para la comparación durante el análisis. Y, finalmente, presenta un bajo registro en credibilidad y exactitud en el tratamiento y enjuiciamiento de los criterios de evaluación (4,4), penalizándole en los que le son de aplicación la falta de respuesta suficiente en términos de eficiencia y sistema de monitoreo del programa.

18,8	**Calidad del informe y utilidad** [Sobre 20]

La evaluación de la calidad y utilidad presenta resultados notables (9,38), pues el informe describe de forma balanceada sus fortalezas y debilidades, diferencia valoraciones descriptivas y valorativas y existe trazabilidad entre hallazgos, conclusiones y recomendaciones. Las recomendaciones que emite son relevantes y, en general, presentan un grado de concreción adecuado, aunque hay excepciones. Además contempla un resumen ejecutivo como parte del informe. Le penaliza el hecho de que alguna de las preguntas formuladas en los TdR no se atienen en esta primera fase de la evaluación.

5,3	**Exhaustividad y adecuación de la documentación** [Sobre 100]

En el criterio de exhaustividad y adecuación de la documentación la evaluación recibe una calificación intermedia (5,2), en tanto que no presenta en anexos los TdR y en estos sólo incluyen algunos de los documentos y datos que especifican el trabajo de campo realizado, aunque si anexa alguno de los instrumentos que aplica. Presenta una estructura adecuada y su lectura está facilitada por gráficos y tablas aunque no con resúmenes de capítulos

10. RESUMEN DE FORTALEZAS Y DEBILIDADES DE LA EVALUACIÓN

FORTALEZAS

- *Desarrolla un meritorio trabajo en la reconstrucción de la Teoría de la intervención siguiendo el Manual de MIDEPLAN.*
- *Formula un completo análisis de los stakeholders.*
- *Contempla un amplio trabajo de campo que abarca una diversidad de perfiles muy adecuada para el objeto que se evalúa.*
- *Ofrece un amplio balance de percepciones de actores clave que puede ser de utilidad para la toma de decisiones*

DEBILIDADES

- *Siendo que la evaluación se define como de ejecución (además de diseño y gestión), no contar datos de monitoreo supone una limitación muy importante. Sólo se trabajaron las bases de datos al efecto de construir la muestra, pero no se explotó para ofrecer información sistematizada sobre las realizaciones físicas y financieras del Beneficio evaluado.*
- *No se presta atención al periodo que se evalúa según los TdR (2014 y primer semestre de 2015), no ofreciendo un solo dato agregado del mismo, sino sólo algún dato aislado de subperiodos*
- *Los TdR presentaron una alta exigencia en cuanto a preguntas (aunque fueron mejoradas de su formulación inicial, según se aprecia en sus anexos), lo que unido a la alta exigencia de productos, genera un encargo de amplio alcance que quizás hubiese requerido una mayor focalización.*
- *Los Anexos presentan información de utilidad pero no tienen un índice que ayude a localizarla, lo cual puede limitar su uso.*
- *En ninguna de las técnicas se contempló si quiera a modo exploratorio la participación directa de la infancia en línea con el enfoque de derechos que promueve esta evaluación. Siendo complejo, alguna muestra en este sentido podría haber supuesto un aporte relevante e innovador, en línea con algunas experiencias en este tipo de programas.*

2.1.6 Evaluación 06: CEN-CINAI

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO

Servicio de Atención y Protección Infantil (API) ejecutado bajo la modalidad de intramuros, perteneciente al Programa CEN CINAI (Centros de Educación y Nutrición y Centros de Atención Integral).

2. INSTITUCIÓN QUE IMPLEMENTA

Dirección Nacional de Centros de Educación y Nutrición y de Centros Infantiles de Atención Integral (Dirección Nacional CEN-CINAI) del Ministerio de Salud.

3. MONTO DEL PROYECTO/ PROGRAMA

El único dato económico hallado en el Informe es el siguiente: "El presupuesto total asignado a la Dirección Nacional de CEN CINAI correspondió en el año 2015 a un poco más de 51 mil millones de colones (51.270.426.969,13)". (p. 30)

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO

El Programa CEN CINAI es una iniciativa que ha venido desarrollando el Ministerio de Salud de Costa Rica desde la década de 1950, en su búsqueda por contribuir a la mejora del estado nutricional y el desarrollo de la población materno-infantil, que vive en condiciones de pobreza y/o riesgo social; mediante los servicios de nutrición preventiva; atención y protección infantil; y promoción del crecimiento y desarrollo; ofrecidos a escala nacional.

En concreto, el Servicio de Atención y Protección Infantil (API), que es el objeto de evaluación, ofrece diariamente a niños y niñas menores de 13 años, de acuerdo a la necesidad de cada familia, los siguientes servicios: a) Servicios de atención integral en un entorno adecuado; b) Promoción del crecimiento y desarrollo infantil; c) Acompañamiento a familias y promoción de servicios comunitarios.

EVALUACIÓN

5. EVALUADORES

Centro de Investigación y Capacitación en Administración Pública (**CICAP**) y la Escuela de Salud Pública (ESP)

Equipo Evaluador: Karol Rojas Araya; Evelyn Zamora Serrano; Cynthia Guzmán Trejos; Ailhyn Bolaños Ulloa; Diana Arce Flores; Esteban Mora Martínez; Víctor Sánchez Brenes.

Equipo Directivo UCR: Mayela Cubillo Mora, Directora CICAP; Mauricio Vargas Fuentes, Director ESP.

Únicamente en el Resumen Ejecutivo que aparece en la Web figuran los nombres concretos del equipo que ha realizado la evaluación. Sin embargo, no se ha hallado en la documentación consultada referencia alguna a los integrantes concretos del Equipo Gestor, del Equipo de Toma de Decisión o del asesor externo, si bien estas figuras aparecían reflejadas en los TdR.

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN	
No especificado en TdR ni en Informe final de evaluación.	**Fecha Informe**:	Diciembre 2015

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL

Evaluación externa (Consultoría nacional – CICAP-UCR) en el marco de la ANE apoyada por FOCEVAL II bajo la modalidad de Padrinazgo (Asesor Internacional) y gestionada mediante la conformación de un Equipo Gestor y un Equipo Tomador de Decisiones (MIDEPLAN- DN CENCINAI)

En los TdR se indica que:

- "Por medio de la orientación metodológica de MIDEPLAN y el Programa FOCEVAL al Ministerio de Salud y puntualmente de la Dirección Nacional de CEN CINAI, se decide evaluar la gestión de los procesos API".
- "se contará con el apoyo del Programa de Fomento de Capacitación en Evaluación (FOCEVAL) a lo largo de todo el proceso de la evaluación, en donde se aportará la colaboración de un experto internacional en la figura de padrino el cual estaría contribuyendo a la calidad del proceso de evaluación, asesorando tanto al Equipo de Gestión de la evaluación, como al Equipo de Evaluadores que la realicen".

Sin embargo, resulta llamativo que a lo largo del Informe de evaluación no se haga ninguna mención explícita a ese apoyo de FOCEVAL.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA:	
72,2	**Calificación global** *[Sobre 100]*
7,5	**Imparcialidad e independencia** *[Sobre 10]*
A lo largo de la evaluación se va poniendo de manifiesto cómo el equipo externo aporta objetividad y rigor en la adopción de diferentes decisiones referidas a la metodología y el trabajo de campo. Se incorporan múltiples perspectivas, pero fundamentalmente a la hora de recoger la información, no tanto en las fases de análisis o presentación de resultados.	
9	**Equidad, participación y transparencia** *[Sobre 10]*
Sobre aspectos de equidad hay poca información (aunque los TdR reflejaban la necesidad de adoptar la perspectiva de género); en cuanto a la participación, estuvo bastante limitada, puesto que muchos aspectos ya venían fijados en el TdR, y la implicación de los actores se tradujo sobre todo como informantes en la recogida de información; en cuanto a transparencia, el balance es positivo, no se observan obstáculos relevantes al respecto.	
33,9	**Credibilidad y exactitud** *[Sobre 50]*
Esta evaluación presenta una situación intermedia en este criterio. Hay algunos aspectos del mismo en los que se producen avances importantes (ej. el objeto de la evaluación y los objetivos que se persiguen están claros, pero hay otros aspectos donde el logro no es total (ej. análisis de grupos de interés; utilización efectiva de la teoría de la intervención). En lo que se refiere a diseño y métodos, como aspecto positivo se adopta una pluralidad de métodos, pero por otro lado no siempre hay una suficiente justificación o explicación sobre el porqué de las decisiones metodológicas adoptadas. Los procedimientos de selección presentan cierto margen de mejora, mientras que en el análisis de datos, los resultados son aceptables, y el abordaje de los criterios es bastante satisfactorio.	
17,8	**Calidad del informe y utilidad** *[Sobre 20]*
Los aspectos incluidos en este criterio obtienen una alta valoración, la vertebración del informe es visible y comprensible, sigue una secuencia lógica y coherente, los elementos descriptivos conducen ordenadamente a los elementos valorativos, y las recomendaciones resultan útiles y suficientemente concretas. El aspecto que presenta mayor debilidad en este criterio es la escasez de explicaciones o aclaraciones acerca del grado de fiabilidad de los datos obtenidos.	
4	**Exhaustividad y adecuación de la documentación** *[Sobre 10]*
Estas cuestiones obtienen una valoración moderada. Los anexos, a pesar de ser extensos y contener múltiple información, no recogen aspectos de especial relevancia, como los TdR o el cronograma y listados de personas o lugares.	

10. RESUMEN DE LA EVALUACIÓN.

FORTALEZAS

- Se describe de manera detallada el objeto de evaluación y la teoría de intervención reconstruida, lo cual ayuda a la contextualización y aporta riqueza al proceso evaluativo puesto en marcha.
- Enfoque mixto cuantitativo/cualitativo, se concede importancia a la complementariedad entre diversas fuentes y técnicas. En el informe se insertan de forma coherente diversas aportaciones cualitativas literales, que enriquecen la lectura.
- La estructura por criterios, que se visibiliza con claridad y vertebra las conclusiones y recomendaciones, permitiendo así partir de las evidencias y hallazgos, pero alcanzando una visión más global.
- Sistemáticamente, cada hallazgo va acompañado del conjunto de evidencias que lo sustentan, y que han sido obtenidas a lo largo del proceso de evaluación y de la aplicación de las diversas técnicas.
- Se produce una considerable reflexión acerca de las limitaciones, implicaciones y retos, así como las buenas prácticas y lecciones aprendidas en el proceso de evaluación.

DEBILIDADES

- Insuficiente desarrollo y justificación del diseño metodológico adoptado (ej. Decisiones, estrategias, técnicas); el Informe entra directa y acríticamente a describir las técnicas utilizadas, sin que se visibilice una reflexión previa sobre su pertinencia para esta evaluación. Tampoco hay evidencias ni se hace explícita en el diseño de la evaluación si se ha incluido, y cómo, la perspectiva de género y de derechos de la niñez y adolescencia
- No hay un abordaje profundo del análisis de datos; fundamentalmente, lo que se señala es que se triangularán los resultados obtenidos por las distintas técnicas aplicadas. Apenas hay alusiones a cómo se ha realizado el análisis cuantitativo (más allá del Anexo con Gráficas y Estadísticas), y en el análisis cualitativo, las estrategias descritas son bastante genéricas.
- No se refleja de forma clara y precisa el número de personas consultadas (total y por cada una de las técnicas).

2.1.7 Evaluación 07: EMPLÉATE

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO
Programa Empléate
2. INSTITUCIÓN QUE IMPLEMENTA
Dirección Nacional de Empleo del Ministerio de Trabajo y Seguridad Social (MTSS).
3. MONTO DEL PROYECTO/ PROGRAMA
Según consta en Cuadro 2, p. 36: Recursos asignados de ¢20.440 millones y Recursos ejecutados de ¢15.483 millones [75,7%]

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO
El Programa Empléate, creado en 2010 y que se integra dentro del Programa Nacional de Empleo (MTSS), constituye una estrategia público-privada de promoción del empleo para personas jóvenes (17-24 años, ampliable a 35 años en caso de discapacidad) en situación de vulnerabilidad (que no estudian ni trabajan, y se ubican por debajo de la línea de pobreza). Esta estrategia se fundamenta en alianzas entre instituciones públicas, sector empresarial y organizaciones no gubernamentales con el propósito de capacitar a esos jóvenes para su inserción al mercado laboral, fortaleciendo su empleabilidad a través de capacitaciones técnicas/ ocupacionales vinculadas con la demanda ocupacional insatisfecha, partiendo de los señalamientos del mercado laboral. A través de este programa se les brinda el apoyo económico necesario, la asesoría y orientación, y la posibilidad de acceder a un proceso formativo.

EVALUACIÓN

5. EVALUADORES		
Equipo Evaluador: Ericka Valerio Mena; Javier Alfaro Araya; Mª Elena Fonseca Quirós; Mario González Zúñiga; Tabatha Carvajal Ruiz. Equipo Tomador de Decisiones: Amparo Pacheco Oreamuno; Andrés Romero Rodríguez; Florita Azofeifa Monge; Sandra Mongalo Chan; Tatiana Ramírez Chavarría. Supervisado por: Victoria Sánchez Esteban (Consultora Internacional) y Gabriela León Segura (Consultora Nacional).		
6. MONTO PARA LA EVALUACIÓN	**7. FECHA DE REALIZACIÓN**	
No especificado en Informe final de evaluación	**Fecha Informe**:	2016

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL
Evaluación externa/mixta en el marco de la ANE apoyada por FOCEVAL II bajo la modalidad de Aprender-haciendo mediante Consultoría nacional–internacional y gestionada mediante la conformación de un Equipo Gestor y un Equipo Tomador de Decisiones. Para la elaboración de esta evaluación se conformaron dos equipos, uno Tomador de Decisiones, y otro equipo Evaluador integrado por representantes del Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) y del Ministerio de Trabajo y Seguridad Social) (MTSS). El proceso evaluativo se desarrolló bajo la modalidad de Aprender Haciendo, con el fin de generar capacidades y conocimientos en el tema de la evaluación, replicables a lo interno de las instituciones, pues el equipo evaluador no estaba compuesto por evaluadores profesionales, sino personas que desean adquirir conocimientos y experiencia en evaluación. Se contó con el apoyo del Programa FOCEVAL como órgano financiador, el cual realizó la contratación de dos consultoras, una internacional y otra nacional, responsables de orientar y asesorar al equipo evaluador. Mediante esta asesoría se brindó capacitación especializada al inicio de cada etapa del proceso (diseño, programación, ejecución y socialización), a través de talleres que aportasen los conocimientos teóricos sobre evaluación combinados con una aplicación práctica a través de esta metodología, basada en la realización de una evaluación completa. En el apartado 3.4., sobre limitaciones del proceso de evaluación, se incluyen valoraciones sobre cómo resultó la ejecución de la evaluación bajo esta modalidad de Aprender Haciendo (p.44), subrayándose como principales limitaciones las siguientes: - Diferentes niveles de experiencia en los integrantes del equipo (respecto a su participación en evaluaciones, así como a la aplicación de técnicas de recolección y análisis de la información). Puesto que esta modalidad de Aprender Haciendo exige que todos los integrantes participen en todas las actividades del proceso, independientemente de su experiencia, eso conduce a una ralentización, dedicando más tiempo a cada una de las etapas de la evaluación. - Se produjo rotación de personal dentro del equipo evaluador, que obligó a generar procesos de inducción a funcionarios que ingresaban en sustitución de otros. - Los integrantes del equipo evaluador tenían en realidad un tiempo limitado para dedicar a la evaluación, ya que generalmente su cargo conllevaba también cumplir con otras actividades. - Se consideró insuficiente el acompañamiento por parte de las consultorías; los evaluadores sintieron la necesidad de contar con un apoyo técnico permanente.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA [Global y por Criterios]:

71,7	**Calificación global [Sobre 100]**
7,5	**Imparcialidad e independencia** [Sobre 100]

Que la propia institución evaluada participara en la realización de la evaluación pudo ser un factor limitante, sin embargo esto se contrarresta con la utilización de una pluralidad de técnicas y fuentes, que ayuda a tener una visión global y completa.

5,5	**Equidad, participación y transparencia** [Sobre 10]

Esta evaluación no destaca especialmente por la implicación y participación de los diversos actores implicados, que son involucrados generalmente desde el rol convencional de suministradores de información (informantes clave). No obstante, en cuanto a transparencia los resultados son satisfactorios, ya que los productos de la evaluación pueden ser consultados por cualquier persona a través de Internet, y las preguntas de evaluación son claras y permiten hacerse una idea adecuada de los focos donde se centra esta evaluación.

34,7	**Credibilidad y exactitud** [Sobre 50]

Las puntuaciones en las que se desglosa este criterio resultan desiguales en esta evaluación. Hay aspectos que están consolidados, como la descripción del objeto de evaluación; la elaboración y utilización de una teoría de la intervención y cadena de resultados; la utilización de métodos mixtos, que cubren tanto lo cuantitativo como lo cualitativo, así como muestras y errores aceptables; también se saca algún provecho de los datos del monitoreo; y se trabaja desde una óptica de criterios.
Como aspectos a mejorar, los objetivos de la evaluación son un tanto genéricos, y se echa en falta el desarrollo de un diseño o modelo de evaluación, puesto que lo que se plasma son fundamentalmente las preguntas de evaluación y los criterios, y las técnicas o fuentes con las que se le da respuesta. Tampoco hay apenas atención al análisis de actores, y todo lo referido a análisis de impactos se plantea como propuesta metodológica a futuro.

17,5	**Calidad del informe y utilidad** [Sobre 20]

Aquí también las puntuaciones en las que se desglosa este criterio resultan desiguales. Como puntos positivos, el Informe da respuesta a las preguntas de evaluación, sigue una secuencia lógica y coherente, presenta un resumen adecuado, y no se entremezcla de forma inadecuada lo descriptivo con lo valorativo. No obstante, como aspectos que se podrían haber enfocado mejor, el informe tiene cierta tendencia a resaltar las debilidades más que las fortalezas del objeto evaluado, y tampoco hay una excesiva atención a las cuestiones de fiabilidad en los hallazgos o conclusiones presentados.

6,5	**Exhaustividad y adecuación de la documentación** [Sobre 10]

Por una parte, hay distintos elementos en el informe que facilitan su lectura, y los anexos contienen información valiosa, si bien se detectan algunas carencias (ausencia de TdR, cronograma de trabajo de campo, o listados detallados de participantes y lugares).

10. RESUMEN DE LA EVALUACIÓN

Esta evaluación se orientó a valorar los niveles de diseño, proceso y resultados del Programa Empléate, con el objetivo de apoyar la toma de decisiones para su mejora. El período abarca los años 2012-2015. En ella se buscó identificar fortalezas y debilidades de la intervención, con el fin de proponer medidas correctivas para mejorar y facilitar la implementación del Programa. Así, el proceso de Evaluación Empléate cumple con varios propósitos: aprendizaje, mejora, transparencia y rendición de cuentas.

FORTALEZAS

- En general, estructura clara, ordenada y coherente, y se incluye información suficiente sobre el objeto de evaluación y sobre el proceso de evaluación desarrollado.
- Enfoque mixto, con técnicas cuantitativas y cualitativas. Diversidad de herramientas (tres encuestas, entrevistas, grupos focales, observación, revisión documental y de expedientes).
- Abundante información, y de contenido relevante. El contenido y la forma en que están redactadas las recomendaciones contribuye positivamente a su puesta en práctica.
- Visión detallada y crítica, aunque eso conduce a que tengan un mayor peso la descripción/valoración de debilidades del Programa que sus fortalezas.

DEBILIDADES

- En base a la matriz diseñada, la evaluación transita directamente desde las preguntas de evaluación a las técnicas utilizadas, sin que se refleje modelo teórico o metodológico alguno que guíe y justifique las diversas decisiones metodológicas adoptadas.
- Escaso protagonismo de los criterios a la hora de estructurar los hallazgos, conclusiones y recomendaciones de la evaluación. Se opta por guiar la estructura según las tres dimensiones de diseño, procesos y resultados.
- Ausencia de un análisis previo de stakeholders; tampoco se celebra ningún taller de inicio con agentes involucrados. Se asume la situación de partida como algo fijo, seguramente porque hay otros aspectos que requieren mayor atención e interés por parte de la evaluación, y por las dificultades para reunir o involucrar a esos agentes. Al hilo de lo anterior, el hecho de que no se contara con la participación (activa y necesaria) del responsable del Programa Empléate, es una limitación que quita fortaleza a la evaluación.
- Aunque los Anexos tienen una extensión considerable dentro del Informe, y presentan información diversa y de utilidad, no van precedidos de un índice ordenado y numerado que ayude a identificar el contenido de cada uno de ellos, lo cual resta calidad y puede suponer un menor aprovechamiento de los mismos por parte de quien realiza la lectura.
- A nivel de formato, hubiera sido deseable una estructuración algo más precisa y clarificada de los diferentes epígrafes y subniveles que componen cada capítulo.

2.1.8 Evaluación 08: OAPVD

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO

Programa de Atención de la Oficina de Atención y Protección a la Víctima del Delito (OAPVD).

2. INSTITUCIÓN QUE IMPLEMENTA

Ministerio Público-Poder Judicial.

3. MONTO DEL PROYECTO/ PROGRAMA

Dentro de los resultados de evaluación, en el bloque sobre recursos financieros, se incluyen datos cuantitativos sobre el presupuesto anual destinado al conjunto de la Oficina, pero no específicamente al Programa de Atención (objeto evaluado). (Gráfico nº7. Historial del presupuesto OAPVD después de la aprobación de la Ley 8720, según año (millones de colones), p. 158).

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO

La evaluación realizada se denomina "Evaluación de la calidad del servicio del Programa de Atención de la Oficina de Atención y Protección a la Víctima del Delito (OAPVD), del Ministerio Público, durante los años 2010, 2011, 2012, 2013" .

La Oficina de Atención y Protección a la Víctima del Delito (OAPVD) se crea en el año 2000. Actualmente cuenta con más de 100 funcionarios/as y se encuentran ubicados en las diferentes provincias del país. Cuenta con dos programas principales, el Programa de Atención y el Programa de Protección. Esta evaluación se ha centrado en el Programa de Atención, que estructura esos procesos de atención en tres fases (inicial, seguimiento y cierre), y en tres áreas de atención (trabajo social, psicología, legal).

EVALUACIÓN

5. EVALUADORES

Se trata de una Investigación evaluativa en el marco de la cooperación académica de la Maestría en Evaluación de Programas y Proyectos de Desarrollo (MEPPD) de la Universidad de Costa Rica y el Centro para Evaluación (CEval) de la Universidad del Sarre, Alemania. Según consta al comienzo del Informe, finalmente la composición del equipo fue la siguiente:

Equipo evaluador: Andrea Meneses Rojas (Coord.); Carolina Zúñiga Zamora; Susanne von Jan; Carmen Marín González; Fainier Candamo Barrantes.

Con la colaboración de:

- CEval: Dr. Prof. Reinhard Stockmann (Director)
- Integrantes de la Comisión de la MEPPD-UCR: M.Sc. Olman Villarreal (Director MEPPD), M.Sc. Nidia Morera (Directora Escuela y Posgrado de Trabajo Social), Dra. Mayela Cubillo (Directora del CICAP).
- Docentes de la Universidad de Costa Rica: M.Sc. Corinna Schopphoff, M.Sc. Ana Victoria Molina, M.Sc. Johanna Alarcón Rivera.

Con el apoyo de:

- Pasantes de la Maestría en Evaluación, Universidad del Sarre: Tania Jensen, Katrin Windolf, Sandra Schopper.
- Investigadores/as de la Universidad del Sarre: Jessica Gödert, Simeón Schächtele, Dietmar Mälzer, Sandra Baethge.
- Asistentes de la MEPPD-Universidad de Costa Rica: Luis Fernando Castillo Vindas, Xinia Mora Valverde.
- Asistentes de CEval-Universidad del Sarre: Lennart Schemeling, Kristin Alle.
- Escuela de Trabajo Social – UCR.
- Centro de Investigación y Capacitación en Administración Pública-UCR.
- Escuela de Administración Pública-UCR.
- Programa Fomento de Capacidades en Evaluación (FOCEVAL): GIZ y DEval.
- Sistema de Estudios de Posgrado - UCR: Dra. Cecilia Díaz Oreiro (Decana).

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN	
No especificado en Informe final de evaluación	**Fecha Informe**:	Junio, 2015

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL

Evaluación académica apoyada por FOCEVAL II bajo modalidad Evaluadores emergentes (con la colaboración de MEPPD-CEval)

Además de estar incluidos en la relación de personas o instituciones involucradas en la evaluación y en los agradecimientos, en diversas páginas se menciona el apoyo de FOCEVAL, por ejemplo ello permitió la contratación de la persona coordinadora.

En el capítulo de lecciones aprendidas y buenas prácticas, se incluyen algunas consideraciones relacionadas con ese apoyo por parte de FOCEVAL, de cara a incorporar mejoras futuras, y relacionadas fundamentalmente con la disponibilidad de tiempo efectivo para participar en las tareas evaluativas:

- "Se debe evitar grupos grandes de trabajo, donde existan más de 5 personas. La presente evaluación, por cuestiones presupuestarias, llegó a contar con un equipo de evaluación de 7 personas con tiempos limitados, además de contar con el apoyo en diferentes tareas y actividades de hasta 10 personas adicionales para permitir el avance de la evaluación. Lo anterior generó el recargo de funciones en algunas personas, e intensificando las labores de comunicación de la coordinadora".
- "Si bien una fortaleza del equipo de evaluación fue la multidisciplinariedad y compromiso de las personas que lo integran, así como su experiencia probada en diferentes áreas de trabajo (trabajo social, evaluación, administración pública, sociología, temas judiciales), los tiempos de dedicación a la evaluación, por falta de recursos, fue muy limitada. Las futuras evaluaciones realizadas desde la MEPPD con la cooperación internacional, deben necesariamente contar con dos personas como mínimo que se dediquen tiempo completo a la evaluación del programa o proyecto".

Sobre la socialización de resultados, se indica que el Informe se difundió en diferentes espacios, por ej. A integrantes del grupo de coordinación de FOCEVAL. Finalmente, señalar que comparativamente el Informe detalla más la relación entre el equipo evaluador y la jefatura y personal de la OAPVD (objeto evaluado), que en la relación entre equipo evaluador y FOCEVAL.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA: [Global y por Criterios]:	
71,7	**CALIFICACIÓN GLOBAL [Sobre 100]**
7,5	**Imparcialidad e independencia** [Sobre 10]
A pesar de que los propios responsables y personal de la OAPVD colaboran en la preparación y realización de la evaluación, se consigue en buena medida los objetivos de imparcialidad e independencia, ya que el equipo evaluador es quien asume las riendas y toma las decisiones metodológicas, además de utilizarse diversas técnicas y fuentes de información.	
8,8	**Equidad, participación y transparencia** [Sobre 10]
No se detectan problemas en plantear una evaluación comprensible y en dar a conocer sus resultados, así como numerosos detalles acerca de su implementación. Sin embargo, como debilidad aparece que apenas hay implicación de otros grupos diferentes a las propias OAPVD.	
39,3	**Credibilidad y exactitud** [Sobre 50]
Un considerable número de ítems englobados en este criterio reciben la máxima puntuación. Puede decirse que esta evaluación presenta calidad metodológica, y tiene en cuenta los diversos aspectos necesarios para aportar credibilidad y exactitud (descripción de objetivos y objeto, análisis de grupos, elaboración de un modelo de evaluación, desarrollo y explicación sobre las técnicas empleadas, etc.) Por la propia naturaleza y temática de esta evaluación, se analizan en mayor medida los efectos (efectos directos) que los impactos propiamente dichos.	
11,8	**Calidad del informe y utilidad** [Sobre 20]
La elaboración del informe resulta cuidada, se intenta ofrecer la mayor información posible, si bien eso mismo supone un elemento de mejora, ya que todo eso no viene acompañado de un esfuerzo de síntesis, no se destacan las ideas más importantes ni aparece un resumen ejecutivo. Por tanto, el lector tiene que realizar un esfuerzo adicional para localizar la información más relevante o de su interés. Es por ello que este criterio recibe la puntuación más baja.	
4,3	**Exhaustividad y adecuación de la documentación** [Sobre 10]
A lo largo del informe y de los anexos se intenta ofrecer la mejor y mayor información posible, no sólo de los resultados de la evaluación, sino de su propio desarrollo metodológico y analítico. Aun así, se echa en falta alguna documentación específica, como el detalle del cronograma o el listado completo de informantes que han participado en la evaluación.	

10. RESUMEN DE FORTALEZAS Y DEBILIDADES DE LA EVALUACIÓN

FORTALEZAS

- La evaluación contiene gran cantidad de información, que una vez procesada y sintetizada por quienes encargaron la evaluación, puede ser de gran interés y utilidad para aplicar al Programa.
- Siendo una evaluación académica, se presta especial atención a elementos conceptuales, aclaraciones terminológicas, alusiones bibliográficas dentro del texto.
- Se refleja de manera expresa el enfoque de evaluación utilizado, que equivale al modelo o teoría de evaluación, y posteriormente el marco evaluativo concreto (diferentes elementos de la matriz).
- Importantes esfuerzos en el desarrollo de análisis de efectos (ej. se elabora un índice de empoderamiento).
- Se vuelcan interesantes reflexiones en el capítulo de lecciones aprendidas y buenas prácticas. En general, a lo largo del Informe aparecen diversas claves sobre el "hacer evaluativo", que resulta de utilidad a los profesionales de la evaluación y a las personas interesadas en conocer con más detalle el proceso de implementación de las evaluaciones.

DEBILIDADES

- En bastantes ocasiones, la extensión en el desarrollo de los contenidos dificulta obtener una idea global; se echa en falta un mayor esfuerzo de síntesis y de resaltar aquello que el equipo evaluador considera más importante. En general, la profundidad y forma como está estructurada la información no facilita la labor de lectura.
- Ausencia de un capítulo específico sobre conclusiones. Éstas aparecen entremezcladas en el capítulo de resultados de la evaluación y en el de recomendaciones. Eso dificulta obtener una visión global y sintética sobre las principales ideas aportadas por la evaluación.
- Elevado número de recomendaciones (que aparecen, además, sin numerar ni priorizar, y se repiten dos veces, ya que aparecen tanto en el capítulo de resultados como en el de recomendaciones). Este hecho puede dificultar su posterior implementación. Además, tienen un nivel de detalle desigual, y en ocasiones no está claro de que se trate de una recomendación propiamente dicha, puesto que se entremezcla con elementos de reflexión en general.
- No hay una participación significativa de las personas usuarias en el diseño o implementación de la evaluación, simplemente adoptan la figura tradicional de "informantes clave".

2.1.9 Evaluación 09: IMAS

PROGRAMA / PROYECTO EVALUADO
1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO
Evaluación de Efectos Generados por los Programas Sociales que ejecuta el IMAS, 2009 [Programa Avancemos, año 2009]
2. INSTITUCIÓN QUE IMPLEMENTA
Instituto Mixto de Ayuda Social. IMAS
3. MONTO DEL PROYECTO/ PROGRAMA
¢23.293.163.400 colones (Cuadro 26. Página 70. Apéndice I)

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO

El Instituto Mixto de Ayuda Social (IMAS), es una Institución autónoma con personería jurídica cuya finalidad es resolver el problema de la pobreza extrema en el país, para lo cual planea, dirige, ejecuta y controla planes nacionales destinados a dicho fin.
Su objetivo fundamental es el de promover condiciones de vida digna y el desarrollo social de las personas, de las familias y de las comunidades en situación de pobreza o riesgo y vulnerabilidad social, con énfasis en pobreza extrema; proporcionándoles oportunidades, servicios y recursos, a partir del conocimiento de las necesidades reales de la población objetivo, con enfoque de derechos, equidad de género y territorialidad; con la participación activa de diferentes actores sociales y con transparencia, espíritu de servicio y solidaridad.
Esta evaluación de los efectos generados por los programas sociales implementados por el IMAS en el año 2009, se dividió en dos subevaluaciones:

- ⇨ Evaluación de Efectos del Programa de Asistencia y Promoción Social, año 2009.
  Con la implementación del Programa de Asistencia y Promoción Social se busca dar apoyo económico para satisfacer necesidades básicas de las familias ubicadas en condiciones de pobreza y vulnerabilidad. El programa está acompañado de acciones de promoción y de orientación individual o familiar que brinden sostenibilidad al impulso inicial, con el objetivo de alcanzar una movilidad social ascendente.
- ⇨ Evaluación de Efectos del Programa Avancemos, año 2009.
  El Programa Avancemos se presenta como un programa de transferencia monetaria condicionada, el cual responde a la Rectoría del Sector Social y Lucha contra la Pobreza, con el objetivo de conseguir la Permanencia en el sistema educativo formal de las y los adolescentes y jóvenes entre los 12 y 25 años en condiciones de pobreza extrema, pobreza y vulnerabilidad.

EVALUACIÓN

5. EVALUADORES

MEPPD Maestría de Evaluación de Programas y Proyectos de Desarrollo. Instituto de Investigaciones en Ciencias Económicas de la Universidad de Costa Rica.

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN			
¢70.000.000 colones	**Fecha Informe**:	30 de Setiembre 2011	**Duración**:	enero-junio 2011

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL

Evaluación académica apoyada por FOCEVAL I bajo la modalidad de Evaluadores emergentes (MEEPD-UCR) y gestionada por IMAS.

El trabajo analizado consiste en la evaluación de los resultados generados a partir de la ejecución de los programas sociales, que ejecuta el Instituto Mixto de Ayuda Social IMAS.
Atendiendo a la documentación revisada en esta meta evaluación no hay indicaciones relativas a la participación de FOCEVAL ni GIZ ni MIDEPLAN en el trabajo realizado, ni a cómo se concretó este apoyo o qué aprendizajes se derivaron del mismo.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA: [Global y por Criterios]:

79,3	**CALIFICACIÓN GLOBAL** [Sobre 100]
10	**Imparcialidad e independencia** [Sobre 100]

La evaluación analizada fue realizada por un equipo técnico externo a la organización que implementa el programa. Asimismo, las fuentes de información seleccionadas garantizaban la inclusión de pluralidad de perspectivas en el proceso evaluativo. Atendiendo a esto, se puede convenir que la evaluación garantizaba su imparcialidad e independencia.

4,6	**Equidad, participación y transparencia** [Sobre 10]

En cuanto a la disponibilidad de documentación hay que poner de manifiesto dos carencias fundamentales:

- La documentación final de la evaluación no se encuentra accesible para el público general, al menos vía web.
- El equipo que ha realizado la meta evaluación no disponía del *Apéndice III: Línea Base ni del Apéndice IV: Algunos aspectos metodológicos para llevar a cabo futuras evaluaciones de efectos*. En relación a este aspecto hay que indicar que el Apéndice III era uno de los productos requeridos en los TdR.

Respecto a la participación, no se desarrollaron actuaciones para implicar a los agentes clave antes de lanzar la evaluación.

41,4	**Credibilidad y exactitud** [Sobre 50]
El diseño metodológico de esta evaluación permite atribuirle un importante grado de exactitud y credibilidad. Si bien es cierto que la planificación previa carecía del establecimiento de una cadena de resultados y de hipótesis de impactos, el uso de diferentes técnicas de análisis cuantitativo y cualitativo, la aleatoriedad de la muestra seleccionada y la triangulación de los resultados obtenidos permiten garantizar el desarrollo de un trabajo evaluativo de calidad.	
17,8	**Calidad del informe y utilidad** [Sobre 20]
El uso de indicadores en el análisis de la información, la presentación de los resultados finales en base a los elementos que componen los programas, y la profusa redacción de conclusiones y recomendaciones, dotan a la documentación final de unos elevados niveles de calidad y utilidad. Asimismo, hay que indicar que la estructuración de la documentación final de la evaluación en un documento metodológico principal (Informe final) y dos apéndices a modo de "subevaluaciones" de programas (Apéndice I relativo al PAPS y Apéndice II relativo al programa Avancemos) permiten un acceso a la información por parte del lector sencillo, agradable y cobre todo práctico. La falta de un resumen estructurado tanto de la evaluación desarrollada, como de la documentación disponible, es lo que impide alcanzar una nota de 10 en este apartado.	
5,5	**Exhaustividad y adecuación de la documentación** [Sobre 10]
Si bien la documentación disponible para el equipo que realiza esta meta evaluación tenía carácter exhaustivo y era adecuada para poder analizarla, se ha revelado como una carencia importante la falta de documentación anexa. En relación a este elemento hay que indicar que no se dispone de información estratégica clave dentro de un proceso de evaluación como puede ser: guías para desarrollar grupos focales, guiones de entrevistas, cronograma de la fase de campo que incluyera listados de lugares visitados y de personas entrevistadas.	

10. RESUMEN DE LA EVALUACIÓN

FORTALEZAS

- El equipo evaluador externo ha garantizado la independencia e imparcialidad del trabajo realizado. Al mismo tiempo, se ha incluido una amplia variedad de perspectivas en el desarrollo del análisis cualitativo, incluyendo tanto a población beneficiaria como a gestores y planificadores de los programas evaluados.
- El uso de técnicas de análisis cuantitativo que contemplan la selección aleatoria de participantes supone contar con la posibilidad de poder inferir los resultados obtenidos con la muestra elegida a la población total beneficiaria. En relación a este aspecto hay que indicar que la muestra poblacional seleccionada destaca por su amplitud y nivel de confianza (más de 2.000 familias beneficiarias analizadas).
- El equipo evaluador ha desarrollado un trabajo de triangulación más que aceptable, obteniendo información de carácter estratégico tanto del ámbito cualitativo como cuantitativo. Al mismo tiempo, la triangulación ha permitido cotejar información obtenida a partir de distintas fuentes, pero relativa a un mismo elemento evaluado.
- El posicionamiento con el que se ha realizado el análisis de los resultados obtenidos ha estado focalizado a la consecución de una valoración de los efectos generados por los programas sociales de IMAS evaluados (PAPS y Programa Avancemos) en la población beneficiaria. Al mismo tiempo, la planificación del proceso evaluativo y su diseño metodológico, han posibilitado la obtención de resultados vinculados a la mejora en la gestión de ambos programas, a su eficacia y a su eficiencia.
- Tanto la decisión de estructurar la documentación final en varios elementos (Informe final, Apéndice I y Apéndice II) como su redacción, dotan a la documentación disponible de esta evaluación de un elevado nivel de calidad y utilidad. Asimismo, este trabajo de evaluación ha respondido adecuadamente a los requerimientos incluidos en los TdR.

DEBILIDADES

- Atendiendo a la situación socio económico de la población beneficiaria de los programas sociales evaluados, se hacía necesario establecer un proceso previo de sensibilización/presentación que facilitara la participación de la muestra seleccionada.
- Durante el desarrollo de esta evaluación se produjo una situación conflictiva entre el diseño teórico inicial de la evaluación y la realidad encontrada en el momento de ejecución de la misma. En este sentido hay que poner de manifiesto que debido a elementos ajenos al equipo evaluador (imposibilidad para localizar a población seleccionada en la muestra, falta de participación de los beneficiarios…) hubo que acometer revisiones metodológicas en el proceso evaluativo tales como cambios en la muestra, manteniendo su carácter aleatorio pero ampliando la selección inicial.
- Se registraron incidencias en las consultas realizadas a la base de datos SIPO. En la mayor parte de los casos se trataba de diferencias existentes entre la información digitalizada y la física a nivel de número de expediente.
- El hecho de efectuar un proceso de selección aleatoria de participantes, disponiendo de una muestra poblacional bastante amplia, y aplicando unos criterios técnicos que minimizan el error permitido y elevan el nivel de confianza de los análisis cuantitativos, permite desarrollar un análisis exhaustivo que posibilita la inferencia de resultados muestrales a la población general. Sin embargo, aun siendo esto así, la no existencia de un grupo de control impide en modo alguno poder inferir los efectos detectados por la evaluación a la exclusiva aplicación de los programas sociales del IMAS analizados.
- Existe una carencia importante en la documentación disponible: no se incluye el Apéndice III relativo al establecimiento de la línea de base (2009). Su importancia es bastante elevada ya que era uno de los productos requeridos en los TdR.

2.1.10 Evaluación 010: PRENATAL

PROGRAMA / PROYECTO EVALUADO
1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO
Atención Prenatal en el Primer Nivel de Atención en la región Brunca y Huetar Caribe
2. INSTITUCIÓN QUE IMPLEMENTA
Ministerio de Salud
3. MONTO DEL PROYECTO/ PROGRAMA
No especificado en el Informe final de evaluación.
4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO
El programa evaluado es el de Atención Prenatal en el Primer Nivel de Atención, que se realizó en Costa Rica en la Región Brunca y la Región Huetar Caribe en el año 2012. Se trata de un programa de atención prenatal que se brinda en los Equipos Básicos de Atención Integral en Salud (EBAIS) vinculado al cumplimiento de la Norma Oficial Atención Integral a la Mujer Durante el Embarazo de Bajo Riesgo Obstétrico. El programa se inserta dentro del Plan Estratégico Nacional para una Maternidad e Infancia Saludable y Segura, 2006-2015 cuyo propósito es "garantizar la salud y el bienestar de las mujeres, las madres, las niñas y los niños, para que tengan oportunidad de una vida saludable y segura, mediante el acceso oportuno y de calidad a salud, educación, trabajo, transporte, en un entorno familiar y social solidario, amigable y seguro". Este Plan contempla los siguientes componentes: 1) atención integral en salud materna, perinatal e infantil, 2) aseguramiento de la calidad, 3) participación y comunicación social, y 4) marco de política pública legal e institucional.

EVALUACIÓN		
5. EVALUADORES		
Ministerio de Salud Caja Costarricense de Seguro Social (CCSS) Asesoría técnica: - Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) - GIZ-FOCEVAL (Asesor etapa de programación)		
6. MONTO PARA LA EVALUACIÓN	**7. FECHA DE REALIZACIÓN**	
No especificado en el Informe final de evaluación.	**Fecha Informe:**	**No se especifica**
8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL		
Evaluación impulsada y realizada por el Ministerio de Salud en coordinación con GIZ-FOCEVAL-MIDEPLAN (conformando un Comité revisor), prestando la GIZ asistencia en su diseño y análisis documental y MIDEPLAN en su diseño, implementación y análisis. Partiendo de que al Ministerio de Salud como ente rector de la Producción Social de la Salud, le compete desarrollar la función estratégica de Evaluación del Impacto de las Acciones en Salud, las autoridades plantearon la necesidad de desarrollar esta evaluación a través de un proceso de coordinación y articulación de la Unidad de Evaluación de Impacto de las Acciones en Salud de la Dirección de Planificación Estratégica de la Salud (Ministerio de Salud), con la Agencia de Cooperación Alemana (GIZ), el Programa de Fomento de las Capacidades en Evaluación (FOCEVAL) y el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN). La GIZ prestó asistencia técnica para la formulación y diseño de esta evaluación que desarrollaba internamente el Ministerio de Salud, conformando un equipo mixto. En el que MIDEPLAN también tuvo una participación relevante, tanto en el Comité revisor, como en la elaboración del diseño de la evaluación y de los instrumentos de recolección de información, en el asesoramiento para aplicación de encuestas, en el proceso de análisis de la información y en la elaboración del informe final. Para el desarrollo del trabajo se apunta que se contó con el Modelo Conceptual y Estratégico de Evaluación de Impacto de Acciones en Salud diseñado por la Unidad de Evaluación de Impacto de Acciones en Salud (UEIAS) del Ministerio de Salud, el cual se indica que está alineado con el Manual Gerencial para el Diseño y Ejecución de Evaluaciones Estratégicas de Gobierno (MIDEPLAN).		

METAEVALUACIÓN	
9. CALIFICACIÓN CUANTITATIVA: [Global y por Criterios]:	
58,2	**CALIFICACIÓN GLOBAL** [Sobre 100]
5	**Imparcialidad e independencia** [Sobre 10]
La selección del equipo evaluador y de los stakeholders, no garantiza en este trabajo unos niveles elevados de imparcialidad e independencia. Se trata de una evaluación mixta, con participación de la institución evaluada en el equipo evaluador y cuyas fuentes seleccionadas no permiten captar una pluralidad de perspectivas. En relación a esto hay que indicar que, si bien inicialmente se contó con la participación de una importante diversidad de actores, tanto durante la realización del trabajo de campo, como en el análisis de información y resultados, el trabajo sólo incluyó a agentes pertenecientes al programa evaluado (técnicos del ministerio de salud, gestores del programa y usuarias del servicio).	
7,8	**Equidad, participación y transparencia** [Sobre 10]
Aunque en la planificación inicial se contempló la participación de grupos de diferentes entidades que aportaron diversidad de perspectivas, esta situación no tuvo continuidad durante los trabajos de evaluación. No existe disponibilidad de toda la documentación. En relación a este aspecto hay que indicar que el informe final no se encuentra accesible para el público general, al menos vía web.	
32,6	**Credibilidad y exactitud** [Sobre 50]

A partir del análisis de la información presentada en el Informe final se puede afirmar que el documento presenta bajos niveles de credibilidad según las preguntas contempladas. Se detallan a continuación varios aspectos que justifican la afirmación realizada:

- Falta información relativa a aspectos económicos del programa y de la propia evaluación.
- No se desarrolla un análisis previo de los agentes informantes clave.
- Aunque en el documento se indica que se va a usar un "Modelo Conceptual y Estratégico de Evaluación de Impacto de Acciones en Salud", el diseño metodológico no incluye elementos esenciales en una evaluación de impacto como puede ser la definición de un contra-factual, el establecimiento de hipótesis previas para dar respuesta a las preguntas de evaluación, o la determinación de diferentes niveles de impacto.
- No se efectúa un proceso de triangulación de información que permita validar los hallazgos registrados en el trabajo de campo.

El informe final no expone de manera concluyente y compresible los logros alcanzados por el programa.

7,5	**Calidad del informe y utilidad** [Sobre 20]

En base a la información presentada en el Documento final, esta evaluación presenta bajos niveles de calidad y utilidad según las preguntas contempladas. Se trata de una evaluación de la calidad de un servicio público que se ha condicionado previamente con la aplicación metodológica de un modelo de evaluación de impacto. Sin embargo, tanto la definición del diseño metodológico como el desarrollo del trabajo de campo no han cumplido con los requisitos mínimos exigibles para una evaluación de impacto.

Asimismo, aunque en el Informe final se parte de una premisa que vincula directamente un nivel de cuidados prenatales aceptables con la aparición de al menos cuatro resultados positivos, no se identifican ni resultados ni impactos del programa. Se trata por tanto de un documento que posee una utilidad limitada, revelándose más como un informe de indicaciones generales para alcanzar unos niveles mínimos de calidad en la implementación del programa de Atención prenatal.

5,3	**Exhaustividad y adecuación de la documentación** [Sobre 10]

De forma global la documentación incluida en anexos es amplia. Sin embargo, se echa en falta el documento de los TdR, así como información relativa al presupuesto (programa y evaluación) y a las entrevistas desarrolladas (Persona entrevistada, lugar, fecha…).

Es necesario indicar también que aunque la forma en la que se estructura el informe a nivel de capítulos es adecuada, la distribución de los epígrafes dentro de los mismos, es manifiestamente mejorable.

10. RESUMEN DE LA EVALUACIÓN.

FORTALEZAS

- El trabajo de planificación incluyó a una importante diversidad de actores que posibilitó contar con variedad de perspectivas.
- La evaluación contó con la participación de población usuaria que aportó un componente cualitativo muy necesario.
- Se contó con la participación de gestores del programa del ámbito local y regional que aportaron una perspectiva estratégica.
- Se contemplaron técnicas de análisis cuantitativas y cualitativas.
- Las conclusiones obtenidas se vinculan con la necesidad de alcanzar unos niveles mínimos de exigencia en la gestión del programa, que garanticen una mejora de la calidad del servicio público evaluado.

DEBILIDADES

Aunque la evaluación adopta particularidades en el campo de la salud y tiene amplia tradición, existen fundamentos teórico-metodológicos que todo proceso evaluativo debe respetar.

- Existen confusiones conceptuales y debilidades muy relevantes en el planteamiento y desarrollo de esta evaluación que merman su calidad y limitan la sistematicidad del enjuiciamiento que debe caracterizar toda evaluación.
  - ⇨ Se plantea un modelo de impactos para evaluar la calidad, cuando son criterios muy diferentes. Es decir, establece previamente el uso de una metodología de evaluación vinculada a medir resultados e impactos cuando el objetivo del trabajo debía ser evaluar la calidad de un servicio público.
  - ⇨ No se explicitan los criterios de valor bajo los que se enjuiciará el programa y se determinará la calidad del servicio.
  - ⇨ No se establecen las variables o componentes en los que se conceptualiza que se desagrega la calidad y que por tanto, debieran guiar la recolección de información y los indicadores.
  - ⇨ No se explicita en qué consiste el modelo metodológico que se aplica.
  - ⇨ No se formulan conclusiones en sentido estricto, sino que se exponen los hallazgos del proceso de manera resumida, si bien el juicio evaluativo que debe caracterizar a las conclusiones, cuando existe, queda diluido en el texto.
  - ⇨ No se establece una estrategia clara de análisis de la información: triangulación, comparación…

Estos motivos hacen complejo aplicarle a este Informe la lista de criterios y preguntas de esta metaevaluación. Otras debilidades son:

- El modelo de cambio que se desarrolla (aunque presentado gráficamente y con poca explicación, parece razonable y un elemento útil), no se usa como eje para articular la evaluación, es decir, no se usa en la evaluación sino que supone un subproducto de la misma.
- Si bien inicialmente se contó con la participación de una amplia variedad de actores de carácter externo al programa, conforme avanzó el proceso de evaluación todas las fuentes de información utilizadas fueron internas (gestores del servicio, expedientes de atención a usuarias, beneficiarias del programa).
- El proceso de evaluación se desarrolló por parte de un equipo evaluador que pertenece al mismo organismo que implementa el programa, lo que garantiza el conocimiento experto pero puede generar mermas en su credibilidad.
- Los resultados obtenidos no incluyen un análisis de la calidad del servicio basado en la comparativa con información externa al programa. Asimismo, el Informe final de evaluación no contempla un análisis de la eficiencia y eficacia del servicio prestado.
- La metodología de evaluación de impacto diseñada no se ha desarrollado de forma correcta. En relación a esto existen importantes carencias relativas al uso de herramientas fundamentales en el desarrollo de una evaluación de impacto (estimación de contra-factual, establecimiento de hipótesis, triangulación de la información recogida,…).

2.1.10 Evaluación 011: PREVENCION CÁNCER (PyA-Cáncer)

PROGRAMA / PROYECTO EVALUADO

1. NOMBRE DEL PROYECTO / PROGRAMA EVALUADO

Servicios de Prevención y Atención del Cáncer en Valle La Estrella, Provincia de Limón, Costa Rica

2. INSTITUCIÓN QUE IMPLEMENTA

Caja Costarricense del Seguro Social

3. MONTO DEL PROYECTO/ PROGRAMA

No se especifica en el informe

4. DESCRIPCIÓN DEL PROYECTO/PROGRAMA EVALUADO

Los servicios de prevención y atención del cáncer que presta la Caja en Valle de la Estrella encuentran su anclaje estratégico en el Plan Nacional para la Prevención y Control del Cáncer 2012 – 2017 (Ministerio de Salud, 2012), cuyo propósito es lograr el control del cáncer mediante un enfoque sistémico, un abordaje integral, una gestión coordinada y la participación de los actores sociales. En este sentido destaca el rol auxiliar de las Juntas de Salud, a través de las que se canaliza dicha participación.

La prestación de estos servicios se articula en dos ámbitos: prevención y atención en tres niveles (primaria, secundaria y terciaria).

Los **servicios de prevención** se prestan en los centros de salud (establecimientos de atención ambulatoria, Equipos Básicos de Atención Integral en Salud (EBAIS) y consultorios de visita periódica). Se cuenta con el apoyo de enlaces de oncología, Asistentes Técnicos de Atención Sanitaria Primaria (ATAP) y otros actores (estudiantes de medicina o trabajadores sociales). Se realizan campañas en distintos medios de difusión, charlas informativas y, ocasionalmente, ferias de la salud.

La **atención primaria** corresponde también a los centros de salud, que ofrecen servicios de consulta externa y detección precoz de enfermedades (entre ellas el cáncer), curaciones y rehabilitaciones de menor complejidad. El personal médico aplica un formulario sobre factores de riesgo y ofrece información; derivando al nivel secundario posibles casos de cáncer y apoya a pacientes y familias.

La **atención secundaria** corresponde a 13 hospitales periféricos y 7 regionales (Tony Facio, en Limón, para los pacientes de Valle de la Estrella). Prestan apoyo al primer nivel, coordinan de enlaces de oncología; ofrecen intervenciones ambulatorias y hospitalarias para especialidades básicas y realizan cirugía poco compleja y pruebas que de ser positivas se derivan al tercer nivel.

La **atención terciaria** corresponde a 3 hospitales nacionales generales y 6 especializados, siendo el Hospital Calderón Guardia, en San José el que atiende a la población de Valle de la Estrella. Presta servicios ambulatorios y de hospitalización más complejos o que requieren alta tecnología y especialización.

EVALUACIÓN

5. EVALUADORES

Equipo de evaluación:
Caja Costarricense de Seguro Social: Karol Aguilar, Virginia Venegas
Juntas de Salud: Matina (Esther Barquero); Siquirres (Norma Barr); Guácimo (Fressy Calderón); Valle de la Estrella (Herman Fernánder; Susana Olivares; Olga Ramírez); Limón (Eugenia Romero).
Equipo de apoyo: Defensoría de los Habitantes (Hannia Silesky; Rodolfo Vargas); MIDEPLAN (Ericka Valerio)
Expertos consultores independientes: Experto evaluación participativa (Esteban Tapella); Facilitadora sesiones de trabajo (Karla Salazar)
Coordinación: DEval (Andrea Meneses; Juan Carlos Sanz)

6. MONTO PARA LA EVALUACIÓN	7. FECHA DE REALIZACIÓN	
No se especifica	**Fecha Informe:**	Junio 2016 - Febrero de 2017

8. DESCRIPCIÓN DE LA MODALIDAD DE IMPLEMENTACIÓN/APOYO DE FOCEVAL

Evaluación participativa impulsada por el Consejo Regional de Juntas de Salud y realizada por un equipo de evaluación con representación de las Juntas de Salud y de la Caja Costarricense del Seguro Social que contó con el apoyo de MIDEPLAN y la Defensoría de Habitantes, la asistencia técnica de consultores en evalaución participativa y facilitación de procesos participativos, y la coordinación de FOCEVAL-DEval.

A partir de la propuesta de la propuesta del Grupo de Coordinación de FOCEVAL de vincular evaluación y participación ciudadana (junio de 2014) se inicia una línea de actividad para promover evalauciones participativas. Esta línea se inicia con un Encuentro con organizaciones sociales para conversar sobre evaluación (julio de 2015). Continúa con una capacitación en evaluación participativa dirigida a estas organizaciones (noviembre de 2015). Y, deriva, en enero de 2016 en la convocatoria de evalauciones participativas dirigidas a estas organizaciones sociales ofreciéndoles apoyo para elaborar sus propuestas. En esta convocatoria fue seleccionada la propuesta presentada por Consejo Regional de Juntas de Salud de la Región Huétar Atlántico, dando inicio en junio de 2016.

Esta modalidad se concreta en apoyar iniciativas en las que los propios stakeholders demandan la evaluación y conforman el equipo evaluador, contando con el apoyo de MIDEPLAN, ofreciendo FOCEVAL asesorías especializadas y asumiendo DEval la coordinación. En esta evaluación el equipo evaluador (Junta de Salud y CCSS) recibió asesoramiento experto (en evaluación participativa y en facilitación de procesos participativos), contó con el apoyo de MIDEPLAN y la Defensoría de Habitantes, así como con la coordinación de FOCEVAL. La modalidad de apoyo estuvo focalizada en desarrollar las capacidades de las Juntas de Salud y de la CCSS para facilitar su apropiación del proceso, garantizando la sistematicidad y el rigor de la evaluación a paritr de los apoyos, asesorías y la coordinación del proceso.

METAEVALUACIÓN

9. CALIFICACIÓN CUANTITATIVA: [Global y por Criterios]:

73,1	**CALIFICACIÓN GLOBAL** [Sobre 100]
5	**Imparcialidad e independencia** [Sobre 10]

La evaluación obtiene una puntuación intermedia en este criterio, al optar por un equipo mixto y recurrir a un número limitado de fuentes que además tuvieron que ser adaptadas para que el equipo evaluador pudiera apropiarse de las mismas.

10	**Equidad, participación y transparencia** [Sobre 10]

Esta evaluación obtiene la mayor puntuación posible en términos de equidad y participación. Se apoya para ello en una implicación sostenida de los stakeholders a lo largo del proceso; se formulan preguntas claras y adaptas precisamente desde la plataforma que constituye el equipo evaluador donde dichos stakeholders están representados; y no sólo hubo talleres de inicio, sino que el propio lanzamiento de la evaluación estuvo precedido de un proceso de desarrollo de capacidades, que se mantuvo a lo largo de la evaluación en las sesiones/talleres de trabajo del equipo evaluador. También en términos de transparencia, la aceesibilidad a través de la web del informe y su resumen ejecutivo supone un factor positivo.

31	**Credibilidad y exactitud** [Sobre 50]

Tomando los subcriterios que vertebran el análisis de credibilidad y exactitud, esta evaluación obtiene buenos registros en cuanto a la delimitación de Metas, objetivos y modelos de cambio (8 sobre 10), también en su Diseño y métodos (8,8 sobre 10) y en la explicitación de sus Procedimientos de selección (10 sobre 10), pues justifica sobradamente su diseño así como las técnicas que aplica y los criterios de selección desplegados en ellas.

Estos buenos registros contrastan con sus logros en cuanto a Análisis de datos (4 sobre 10), donde le penaliza el hecho de que sea exclusivamente cualitativa no triangule datos cuantitativos con ellos, no contemple datos del monitoreo y hace un uso muy limitado de fuentes externas. Asimismo, sus registros en cuanto al tratamiento de los Criterios de evaluación (5,2 sobre 10) está derivado de que si bien responde a todas las preguntas de evaluación a partir de los hallazgos, el juicio evaluativo queda diluido en una suerte de síntesis de hallazgos.

18,8	**Calidad del informe y utilidad** [Sobre 20]

El informe obtiene una muy valoración sobre su calidad y utilidad, al dar respuesta suficiente a las preguntas de evaluación, presentando de manera balanceada puntos fuertes y débiles, y existiendo una trazabilidad adecuada entre hallazgos, conclusiones y recomendaciones.

8,3	**Exhaustividad y adecuación de la documentación** [Sobre 10]

En cuanto a la documenta que aporta en sus anexos, contempla la matriz de evaluación, las guias de los instrumentos aplicados y las fuentes consultadas, así como los principales datos del cronograma del trabajo de campo (que también aparecen a lo largo del informe), aunque no se detallan las personas participantes o las organizaciones a las que pertencen en todos los casos, siendo esto especialmente necesario en las entrevistas para poder valorar su cobertura de perfiles específicos.

10. RESUMEN DE LA EVALUACIÓN.

FORTALEZAS

- El desarrollo de capacidades en evaluación en el seno del equipo evaluador conformado por las Juntas de Salud y la CCSS.
- La utilidad potencial del informe para las Juntas de Salud y la CCSS, en cuanto a su capacidad para inducir mejoras y la factibilidad de aplicar las recomendaciones por el hecho de apreciarse un importante nivel de apropiación del proceso.
- La sistematicidad y el rigor del planteamiento evaluativo, en el que no se aprecia que el conocimiento experto y directo del equipo evalaudor interno (Juntas de Salud y CCSS) haya sido un componente clave para fundamentar conclusiones no basadas en los hallazgos del proceso, lo cual siempre constituye un riesgo cuando se hace una evaluación interna (pre-juicio).
- La prudencia de las preguntas de evaluación, en las que se aprecia que ante el reconocimiento de las limitaciones del diseño evaluativo, se ha desplegado un proceso dinámico de focalización y priorización, ajustando expectativas.

DEBILIDADES

- Los métodos y técnicas cualitativas encuentran su representatividad cuando alcanzan la saturación de los discursos recabados, sin embargo en este informe de evaluación no se aporta información suficiente para juzgar este aspecto clave.
  - Esto es así especialmente en las entrevistas y grupos focales, donde tampoco se aprecia en qué medida abarcan cada uno de los perfiles relevantes en los diferentes niveles y ámbitos de la prevención y la atención del cáncer, tanto en la vertiente de prestación directa del servicio en lo local como en cuanto al componente institucional desde la perspectiva de la CCSS.
  - Esto también aplica al juego sobre mitos del cáncer, que se aprecia más como subproducto del proceso (útil a futuro) que como soporte de hallazgos que faciliten juicios concluyentes sobre grado de conocimiento, por su limitada aplicación.
- El componente valorativo de la evaluación, centrando en la formulación del juicio de valor que debe caracterizar a las conclusiones, ha quedado diluido en este informe, en el que suele recurrirse a un balance-síntesis de las percepciones recabadas, especialmente en aquellos criterios en los que se aprecia una menor vertebración a partir de sus principales dimensiones (eficacia y eficiencia frente a pertinencia).
- El escaso uso de fuentes externas e información del monitoreo que aparece condicionado por la premisa de la evaluación participativa sobre que sea el propio equipo evaluador (Juntas de Salud y CCSS) el que procese la información y tome las decisiones, lo que exige acotar, derivadando en la renuncia a información muy relevante para la solvencia de los hallazgos.
- La orientación de la evaluación hacia el desarrollo de capacidades de un equipo evaluador interno que conoce profundamente el contexto y los servicios evaluados, ha derivado en una menor precoupación en el informe final por ofrecer información relevante para juzgar la solvencia de los hallazgos.

2.2. Hallazgos: apreciación general de las evaluaciones

Tras presentar en el epígrafe 2.1 la ficha sintética de las 11 evaluaciones objeto de metaevaluación con los resultados cuantitativos y cualitativos de cada una de ellas, en este apartado se presenta en primer lugar una sinopsis de las calificaciones de las evaluaciones, con visión de conjunto y, en segundo lugar, un análisis de los patrones de comportamiento que se aprecian en las evaluaciones respecto a los criterios y dimensiones contemplados en esta metaevaluación para enjuiciar su calidad.

2.2.1. Sinopsis de la calificación de las evaluaciones analizadas por la metaevaluación.

En la siguiente tabla se presenta la calificación obtenida por las 10 evaluaciones, indicando también la ponderación de cada criterio/subcriterio en la puntuación total (100 puntos), así como la calificación promedio. Se juzga así, la calidad metodológica, la validez de los resultados y los términos de referencia los criterios preestablecidos, que corresponden a estándares internacionales definidos por FOCEVAL.

Tabla 5. Sinopsis de calificaciones de las evaluaciones según criterios aplicados y dimensiones

ESQUEMA DE META-EVALUACIÓN		EVALUACIONES											
Criterios	Subcriterios / Dimensiones *Ponderación*	PROMEDIO	DRAT	EDU-INCL	Germina-dora	PND	RED CUDI	CEN-CINAI	EMPLEATE	OAPVD	IMAS	PRE-NATAL	PyA CÁNCER
			1	2	3	4	5	6	7	8	9	10	11
1. Imparcialidad e independencia:	*10 puntos*	**8,0**	10	7,5	10	10	7,5	7,5	7,5	7,5	10	5	5,0
2. Equidad, participación y transparencia:	*10 puntos*	**7,4**	5,5	10	7,5	6,5	6	9	5,5	8,8	4,6	7,8	10,0
3. Credibilidad y exactitud:	*50 puntos*	**35,1**	46,1	27,3	37,5	32,5	30,0	33,9	34,7	39,3	41,4	32,6	31,0
3.1	Metas, objetos y modelo de cambio *[10p.]*	**6,9**	9,3	5,1	9,3	5,3	7,1	6,4	6,0	8,6	5,4	5,3	8,0
3.2	Diseño y métodos *[10 p.]*	**7,8**	8,2	7,2	8,0	8,8	7,2	6,2	6,2	10,0	8,2	6,5	8,8
3.3	Procedimiento de selección . *[10p.]*	**6,6**	10,0	5,0	5,0	5,0	5,0	5,0	7,5	5,5	10,0	10,0	5,0
3.4	Análisis de datos *[10p.]*	**6,2**	10,0	3,0	6,3	5,5	5,3	6,3	7,5	7,5	7,8	5,5	4,0
3.5	Criterios de evaluación *[10p.]*	**7,6**	8,6	7,0	9,0	8,0	5,4	10,0	7,5	7,8	10,0	5,3	5,2
4. Calidad del informe y utilidad:	*20 puntos*	**16,4**	18,8	18,8	16,5	18,8	18,8	17,8	17,5	11,8	17,8	7,5	18,8
5. Exhaustividad y adecuación de la documentación	*10puntos*	**5,5**	6,5	5,5	6,5	5,3	5,3	4,0	6,5	4,3	5,5	5,3	8,3
	TOTAL	**71,3**	**86,9**	**69,1**	**78,0**	**73,1**	**67,6**	**72,2**	**71,7**	**71,7**	**79,3**	**58,2**	**73,1**

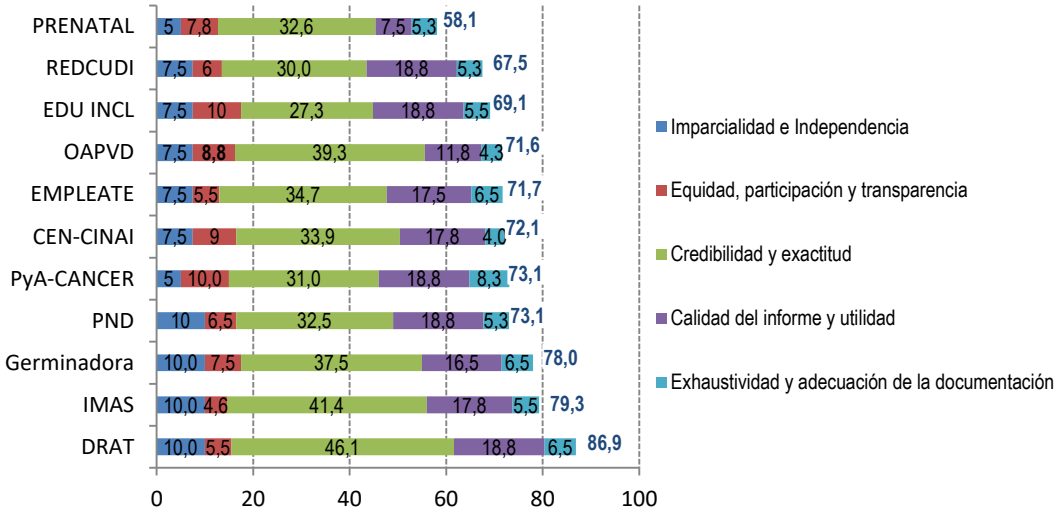
Fuente: elaboración propia a partir de Fichas de registro de evaluaciones

- Las puntuaciones de las evaluaciones se ubican entre los 86,9 puntos de DRAT y los 58,2 de PRENATAL.
- En Imparcialidad de independencia destacan DRAT, Germinadora, PND e IMAS.
- En Equidad, participación y transparencia destacan PyA-CÁNCER y EDU-INCL, seguidos de CEN-CINAI y OAPVD.
- En Credibilidad y exactitud destacan DRAT, IMAS, OAPVD y Germinadora.
  - En la delimitación de Metas, objetos y modelo de cambio destacan DRAT, Germinadora, OAPVD y PyA-CÁNCER, seguidos de REDCUDI.
  - En Diseño y métodos destacan OAPVD, PND y PyA-CÁNCER seguidos de DRAT e IMAS y Germinadora.
  - En la explicitación y adecuación de los Procedimientos de selección destacan DRAT, IMAS y PRENATAL.
  - En la Idoneidad del Análisis de datos destaca DRAT e IMAS seguidos de EMPLEATE y OAPVD.
  - En el tratamiento y respuesta a los Criterios de evaluación que les son de aplicación destacan IMAS y CENCINAI seguidos de Germinadora y DRAT.
- En Calidad del informe y utilidad destacan DRAT, EDU-INCL, PND, REDCUDI y PyA-CÁNCER.
- En Exhaustividad y adecuación de la documentación que acompaña al informe únicamente destaca PyA-CÁNCER, siendo las siguientes mejor posicionadas DRAT, EDU-INCL, Germinadora y EMPLEATE.

2.2.1. Descripción de patrones y tendencias generales

A la vista de los resultados obtenidos al aplicar los criterios y preguntas de esta metaevaluación, es posible apreciar una serie de patrones de comportamiento de las evaluaciones.

Gráfico 1. Puntuaciones de las evaluaciones según criterios aplicados



Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

En cuanto a **imparcialidad e independencia** la puntuación promedio es 8 sobre 10, marcando la pauta una evaluación externa (55%)[11], en la que las fuentes de información se eligieron de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados (64%)[12].

Tabla 6. Puntuaciones de las evaluaciones en el criterio de Imparcialidad e independencia

		1 DRAT	2 EDU INCL	3 Germinadora	4 PND	5 REDCUDI	6 CEN-CINAI	7 EMPLEATE	8 OAPVD	9 IMAS	10 PRENATAL	11 PyA-CANCER
1.	**Imparcialidad e Independencia**	*10*	*7,5*	*10*	*10*	*7,5*	*7,5*	*7,5*	*7,5*	*10*	*5*	*5*
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	10	5	10	10	10	10	5	5	10	5	5
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados?	10	10	10	10	5	5	10	10	10	5	5

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

En este punto cabe señalar que las diferentes modalidades de apoyo ofrecidas por FOCEVAL han conllevado la configuración de equipos de evaluación en los que a partir de la lectura de los informes no resulta inmediato adivinar si estamos ante una evaluación externa o mixta[13]. Esto es así por la pluralidad de entidades

[11] El 45% restante fueron evaluaciones mixtas, o al menos así fueron calificadas al no quedar claro en los informes la participación de la entidad responsable del programa/proyecto evaluado.

[12] El 36% restante se trata de evaluaciones en las que las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas, no encontrándose ninguna en la que se aprecie que no se ha logrado capturar esta pluralidad de perspectivas en ninguna etapa del proceso.

[13] Es ilustrativo el caso dela evaluación 02, donde expresamente se indica que se trata de una evaluación externa pero se incluye la institución responsable del programa (MEP) en el equipo evaluador y el propio informe señala como debilidad su participación en la evaluación, no quedando claro si participó como gestora de la evaluación o cómo parte de facto del equipo evaluador.

implicadas en los procesos (normalmente además de la persona o entidad externa que realiza/coordina la evaluación, se cuenta con la participación de MIDEPLAN en el proceso como órgano gestor de la evaluación –en alguna ocasión como asistencia técnica en el diseño-, y la propia administración responsable del programa o proyecto evaluado), y la escasa atención que prestaron los Informes a delimitar el rol jugado por cada uno de ellos (asistencia técnica, enlace institucional para facilitar el acceso a la información o informantes, gestor de la evaluación).

La imparcialidad e independencia ha sido precisamente una de las alusiones que han realizado los equipos de evaluación en las pocas reflexiones observadas en los informes sobre las modalidades de apoyo de FOCEVAL. Concretamente, se ha apuntado que las evaluaciones bajo la modalidad de aprender-haciendo deben ser cuidadosamente planteadas para que no se vea afectada la independencia de la persona o equipo evaluador externo, clarificando roles y funciones a lo largo del proceso.

> "En la planificación, la organización y la ejecución de procesos de evaluaciones públicas, la posibilidad de participación de miembros de la unidad ejecutora de la intervención que genera una serie de condiciones que deberán ser analizadas a priori: función que cumplen en el equipo evaluador, actividades que técnicamente están habilitadas para realizar, compromisos y productos de su participación en la evaluación. Además, es fundamental establecer en los procesos de negociación interinstitucional, los elementos de carácter ético y conceptual que implica integrar personal que participó de la ejecución de la intervención". [Evaluación 02 EDU-INCL, Pág. 16].

En cuanto a la **equidad, participación y transparencia**, la puntuación promedio es 7,4 sobre 10, destacando sobremanera la evaluaciones que prestaron mayor atención a la participación (PyA-CÁNCER y EDU-INCL).

El único aspecto en el que se aprecia un patrón de comportamiento homogéneo es la Claridad y adaptación al contexto de las preguntas de evaluación (91%), donde tan sólo la evaluación de REDCUDI mostró que siendo claras las preguntas, su formulación era compleja al incluir cada pregunta varios criterios de evaluación, generando un encargo de amplio alcance. Cabe apuntar a este respecto que tan sólo se ha dispuesto de los TdR de tres evaluaciones (27%), aunque es generalizado en los informes reproducir las preguntas de evaluación que aquellos contenían. No obstante, no está extendida ni en los TdR ni en los informes la explicación del proceso de detección, priorización y formulación de las preguntas de evaluación.

Tabla 7. Puntuaciones de las evaluaciones en el criterio de Equidad, participación y transparencia

		1 DRAT	2 EDU INCL	3 Germinadora	4 PND	5 REDCUDI	6 CEN-CINAI	7 EMPLEATE	8 OAPVD	9 IMAS	10 PRENATAL	11 PyA-Cancer
2.	**Equidad, participación y transparencia**	***5,5***	***10***	***7,5***	***6,5***	***6***	***9***	***5,5***	***8,8***	***4,6***	***7,8***	***10***
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos?	1	10	10	5	5	5	1	5	1	10	10
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	10	10	10	10	5	10	10	10	10	10	10
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	0	0	0	0	5	10	0	0	10	0	0
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	1	10	5	1	5	10	1	10	1	10	10
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	10	10	5	10	10	10	10	10	1	1	10

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

En cuanto a los TdR analizados en dos casos se aprecia una formulación clara y comprensible (CENCINAI e IMAS), destacando especialmente en este sentido los de CENCINAI. Los TdR de REDCUDI conforman un documento extenso que presenta una estructura que trata de exponer con claridad el complejo entramado institucional y programático en el que se inserta el Beneficio que se evalúa y su redacción es comprensible. Dicho esto, no es fácil para el lector formarse una idea intuitiva del objeto que se evalúa, pues los propios términos o conceptos clave (como Beneficio, Alternativa o Modalidad) hubiesen precisado una mayor explicación y claridad. En cuanto a la estructura, puede considerarse adecuada en términos generales si bien su extensión hubiese precisado un índice de contenidos y de anexos que facilitase su uso y lectura.

En cuanto a la equidad y la participación en cinco evaluaciones (45%) se indica expresamente la realización de un taller de inicio orientado a presentar la evaluación y facilitar tanto la participación en general como el acceso a la información e informantes en particular, dirigiéndose normalmente a los principales stakeholders pero no a los grupos meta de los proyectos. En algunos de estos casos, la finalidad de este taller estuvo asociada más que a la presentación global de la evaluación, a la construcción participada de la teoría de la

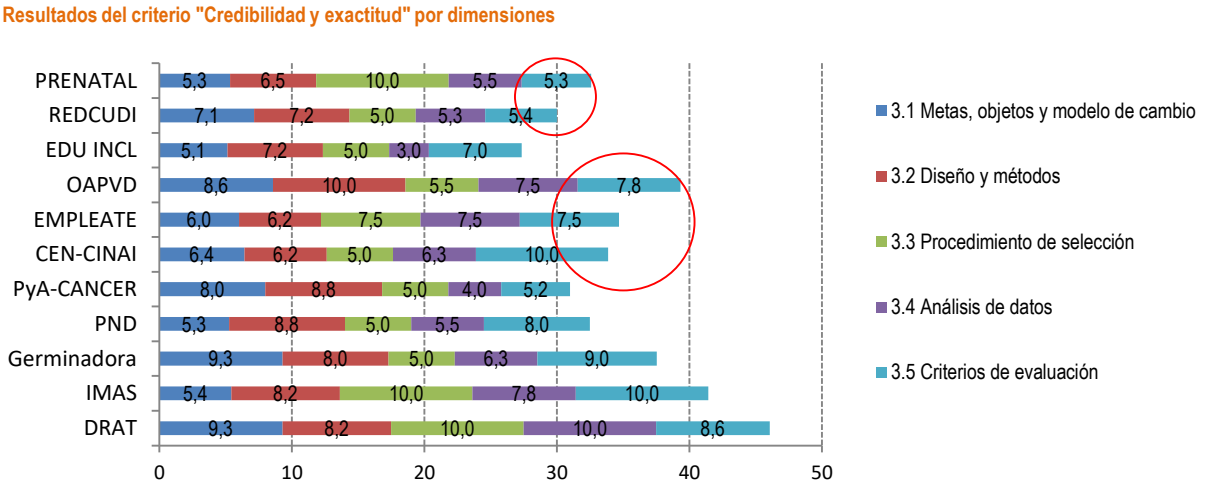
intervención o a contrastar la propuesta inicial diseñada por los equipos evaluadores, realizándose por tanto en los primeros momentos del proceso evaluativo o durante su implementación. En otras dos evaluaciones se aprecian otros mecanismos para facilitar esta implicación, normalmente a través de entrevistas específicas o reuniones con una muestra de actores clave. Y, en el resto de evaluaciones (36%: IMAS, EMPLEATE, PND y DRAT) no se celebraron ni estos talleres de inicio ni se aprecia otras alternativas para facilitar la participación en sus inicios.

La involucración activa de los grupos meta y otros grupos interesados en la planificación y la implementación de los procesos de evaluación presenta un comportamiento heterogéneo, destacando en un extremo tres evaluaciones en las que no se aprecia participación (DRAT, EMPLÉATE e IMAS) y, en el otro extremo, cuatro evaluaciones con participación activa en todas sus etapas (PyA-Cáncer, EDU-INCL, Germinadora y PRENATAL). En el resto de evaluaciones (36%), encontramos o bien una participación más centrada sólo en la planificación o más habitualmente sólo en la implementación; o bien que esta participación adquiere un rol tradicional en el que los grupos meta o stakeholder adquieren la condición de informantes clave a los que se le hacen consultas (pasivos), siendo más bien evaluaciones consultivas que evaluaciones participativas genuinas.

Y, en términos de transparencia, está generalizado el acceso a los informes de evaluación a través de internet desde la web de MIDEPLAN (82% de las evaluaciones), siendo tan sólo 2 evaluaciones (IMAS y PRENATAL) para las que no se localizó dicho informe en internet.

En cuanto a la **Credibilidad y exactitud de las evaluaciones**, el análisis de los patrones de comportamiento se realiza a partir de las cinco dimensiones en las que se concreta este criterio de metaevaluación. El Gráfico 2, en el que las evaluaciones aparecen de abajo hacia arriba de mayor a mejor puntuación total, se aprecia que algunas de las que ocupan posiciones intermedias alcanzan mejores resultados en términos de credibilidad y exactitud, destacando OAPVD, EMPLÉATE y CEN-CINAI, pero también PRENATAL y REDCUDI.

Gráfico 2. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud según dimensiones



Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

En cuanto a la definición y delimitación de las **Metas, objetos y modelos de cambio**, la pauta la marca un informe de evaluación en el que se describe su meta y objetivos –general y específicos- (100%), y se delimita el objeto evaluado (82% de manera exhaustiva y 18% con algunas imprecisiones).

En esta delimitación del objeto un aspecto que tiene escasa presencia en los informes es la dimensión financiera del programa evaluado (volumen e importes por años), algo que se traslada a la implementación y datos de seguimiento. Asimismo, el 64% de las evaluaciones contempló en su diseño un análisis de los stakeholders (36% no lo contempla) y un análisis del contexto evaluativo (100%; aunque en el 82% de casos éste o no es exhaustivo o no se aprecia un vínculo expreso con la teoría de la intervención que se evalúa). Y, en el 82% de las evaluaciones se contempló el diseño de una teoría de la intervención, cadena de resultados o lógica de impacto (en el caso del PND no aplica este ítem y en el caso del IMAS no se aprecia su uso).

Tabla 8. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Metas, objetivos y modelo de cambio

		DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA-CANCER
3.1	**Metas, objetos y modelo de cambio**	*9,3*	*5,1*	*9,3*	*5,3*	*7,1*	*6,4*	*6,0*	*8,6*	*5,4*	*5,3*	*8,0*
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	10	5	10	10	10	10	5	10	10	10	10
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	10	5	10	5	10	10	10	10	10	10	10
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	10	10	10	1	10	5	1	10	1	1	5
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	10	5	10	0	5	5	10	10	1	5	10
a	¿Se formularon hipótesis de impacto concluyentes?	10	5	5	0	5	5	0	10	1	0	0
b	¿Se diferencia entre los diferentes niveles de impacto?	10	1	10	0	5	5	5	5	5	1	0
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	5	5	10	5	5	5	5	5	10	5	5

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

Cabe apuntar que si bien está generalizada la construcción de una teoría de la intervención, es también frecuente que no se aprecie su uso como eje vertebrador del proceso evaluativo sino que se convierte en un subproducto del proceso evaluativo, relevante, pero accesorio y no integrado.También es generalizado una cierta diferenciación de niveles de impacto a nivel de la construcción de las cadenas de resultados (80%); igual que ocurre con las hipótesis de impacto, cuya formulación suele estar implícita, y en pocas ocasiones es explícita o se aprecia que sean concluyentes (DRAT y OAPVD).

En cuanto al **Diseño y métodos**, no se detectan grandes deficiencias en la adecuación de los diseños de las evaluaciones a los intereses de las mismas (entendidos como sus objetivos y preguntas), pues en ninguna evaluación se aprecia desajuste y sólo en 3 se detectan algunas limitaciones en el diseño (integración de la perspectiva de género o enfoque de derechos/infancia) o imprecisiones derivadas de unos objetivos o preguntas de evaluación relativamente genéricas o poco especificadas.

Tabla 9. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Diseño y métodos

		1	2	3	4	5	6	7	8	9	10	11
		DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA-CANCER
3.2	**Diseño y métodos**	*8,2*	*7,2*	*8,0*	*8,8*	*7,2*	*6,2*	*6,2*	*10,0*	*8,2*	*6,5*	*8,8*
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	10	10	10	10	10	5	5	10	10	5	10
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	10	5	5	5	5	5	1	10	10	5	10
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	10	1	5	0	1	1	5	10	1	1	0
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	10	10	10	10	10	10	10	10	10	10	5
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)?	1	10	10	10	10	10	10	10	10	5	10

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

Dicho esto, la pauta en las evaluaciones es un escaso peso de la reflexión metodológica de los equipos evaluación en el diseño, planteando alternativas o justificando la opción elegida, algo que se traslada a la definición y uso de enfoques o modelos teóricos de evaluación. En estos casos lo habitual es que en los capítulos de diseño o metodología se entre de lleno en la especificación de los contenidos de la matriz de evaluación o en el abordaje de técnicas de investigación que se plantean para obtener respuestas a cada pregunta evaluativa. Concretamente la justificación de las técnicas elegidas es habitual (sólo en una evaluación no se aprecia, EMPLÉATE), en 3 evaluaciones se explicitan con suficiencia (OAPVD, IMAS y DRAT), y en el 64% restante esta justificación existe aunque se califica como insuficiente.

En los diseños no es habitual que se contemple lo contrafactual, solo 2 evaluaciones lo contemplaron especificando cómo (DRAT y OAPVD), y en otras 2 (Germinadora y EMPLEATE) se consideró alguna aproximación (grupos de comparación) o bien se contempló pero no se especificó suficientemente[14].

La pauta de comportamiento más homogénea en el diseño es, sin duda, el uso de diversos métodos de levantamiento de datos (82%)[15]que cubren aspectos cuantitativos y cualitativos (91%), siendo la excepción PyA-Cáncer, en la que sólo se usan métodos cualitativos.

En cuanto a los **Procedimientos y criterios de selección** el patrón de comportamiento es que las evaluaciones centren sus esfuerzos en detallar los que aplican en las encuestas, en coherencia con la importancia que suelen otorgar a esta técnica (en línea con la que muestran TdR, algo especialmente significativo en los de REDCUDI); siendo menos habitual que los especifiquen en el caso de las técnicas cualitativas[16].

Tabla 10. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Procedimiento de selección

		1 DRAT	2 EDU INCL	3 Germinadora	4 PND	5 REDCUDI	6 CEN-CINAI	7 EMPLEATE	8 OAPVD	9 IMAS	10 PRENATAL	11 PyA-CANCER
3.3	Procedimiento de selección	*10,0*	*5,0*	*5,0*	*5,0*	*5,0*	*5,0*	*7,5*	*5,5*	*10,0*	*10,0*	*5,0*
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	10	5	5	5	5	5	5	10	10	10	5
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? Se entiende que aplica esta pregunta para las Encuestas	0	5	5	5	5	5	10	1	10	10	0

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

Para ponderar esta situación puede observarse que si en el 100% de las evaluaciones se especifican los procedimientos y criterios de selección, en el 64% se ha valorado que sólo las especifican en algunos casos, siendo precisamente las técnicas cualitativas las que destacan en este sentido, aunque en algún caso también se trata de falta de información relevante sobre la encuestación. Y, en cuanto al manejo de técnicas de investigación lo habitual es contemplar alguna encuesta (82%; las excepciones son DRAT y PyA-Cáncer), si bien en 5 de estas evaluaciones obtienen una puntuación intermedia en la idoneidad de las muestras, normalmente el motivo es que los errores que aceptan se consideran elevados[17], aunque también hay algún caso en que esta valoración deriva de falta de información para su valoración (PND).

En cuanto al **Análisis de datos**, la pauta la marca una evaluación en la que se utiliza la triangulación (91%; la excepción es PRENATAL, y destaca la triangulación efectuada en PND)[18]; se cuantifican los datos relativos a respuestas cuando es posible (91%); y se usan datos del monitoreo (73%; y destaca el caso de EMPLÉATE)[19] aunque en ocasiones de manera insuficiente (36%)[20]. Y en cuanto al recurso de fuentes externas de datos para el análisis es algo que si bien realiza el 64% de las evaluaciones, cabe apuntar que es más habitual que se utilice en mayor medida para la contextualización de las intervenciones que para el análisis en sí, mostrándose pues como un recurso poco explotado por las evaluaciones para fortalecer los hallazgos y facilitar la comparación.

	1	2	3	4	5	6	7	8	9	10	11

[14] En la evaluación del PND se consideró que esta pregunta no aplicaba, al tratarse de una evaluación del diseño metodológico del PND.

[15] Salvo DRAT que sólo usa 1 –entrevistas- y PRENATAL que sólo usa 2 –entrevistas y grupos focales-.

[16] Esto se muestra como un déficit especialmente significativo cuando se recurre a estudios de casos, pues estos se configuran como una amalgama de otras técnicas de investigación, en los que esta ausencia se hace más visible.

[17] Estos son los casos de EDU-INCL, Germinadora, REDCUDI y CENCINAI.

[18] En algunos casos más que una triangulación en sentido estricto (obtener una misma información a partir de diferentes fuentes) lo que se hace es que se utiliza una diversidad de técnicas pero no queda del todo claro en qué medida estas diversas técnicas incorporan las mismas o similares preguntas sobre determinados aspectos como base para la triangulación, por lo que sería más correcto decir que aplican una diversidad de perspectivas y usan la contrastación. El caso de la Evaluación del PND constituye un buen uso de la triangulación en el análisis, claro y sencillo.

[19] Las excepciones son EDU-INCL y REDCUDI.

[20] Son los casos de Germinadora, CENCINAI, OAPVD y PyA-Cáncer.

Tabla 11. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Análisis de datos

		DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA-CANCER
3.4	Análisis de datos	*10,0*	*3,0*	*6,3*	*5,5*	*5,3*	*6,3*	*7,5*	*7,5*	*7,8*	*5,5*	*4,0*
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	10	5	10	10	5	10	5	10	10	1	5
3.4.2	¿Se cuantifican las respuestas siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	10	5	5	1	10	5	5	10	10	10	5
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	10	1	5	10	1	5	10	5	10	10	1
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	10	1	5	1	5	5	10	5	1	1	5

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

A este respecto, es preciso apuntar que los TdR que se han podido consultar en esta metaevaluación, tampoco suelen especificar la importancia o la necesidad de contemplar estos recursos (datos del monitoreo y fuentes de datos externos), siendo especialmente llamativa en el caso de los datos del monitoreo, por cuanto que en ocasiones se trata de evaluaciones intermedias (Germinadora) o evaluaciones que juzgan la ejecución del programa o proyecto (REDCUDI).

En cuanto al tratamiento adecuado y la respuesta concluyente a los **Criterios de evaluación**, se obtiene una alta puntuación promedio (7,6 sobre 10), y se puede afirmar que la pauta de comportamiento la define una evaluación que contempla la relevancia/pertinencia (91%), la eficacia -logro de metas definidas- (82%), que se pronuncia -implícita o explícitamente- sobre la existencia y uso de un sistema de monitoreo en los programas/proyectos, y que trata de identificar cambios observados y reconducirlos al proyecto, aunque el carácter concluyente con que lo hagan sea variable.

Tabla 12. Puntuaciones de las evaluaciones en el criterio de Credibilidad y exactitud: dimensión Criterios de evaluación

		1	2	3	4	5	6	7	8	9	10	11
		DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA-CANCER
3.5	Criterios de evaluación	*8,6*	*7,0*	*9,0*	*8,0*	*5,4*	*10,0*	*7,5*	*7,8*	*10,0*	*5,3*	*5,2*
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa											
a)	es relevante/pertinente para el país, la región y el grupo meta	10	5	10	10	10	10	10	0	10	5	10
b)	logró las metas definidas	10	10	10	10	10	10	10	0	10	5	5
c)	es eficiente	5	0	10	0	1	10	5	10	10	0	5
d)	causó impactos (outcome e impacto) deseados y no deseados.	10	0	0	0	0	0	5	10	10	0	0
e)	causó impactos duraderos	10	5	0	5	0	0	0	0	0	0	0
f)	cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	10	10	10	10	1	10	10	1	10	10	1
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	5	5	5	5	5	0	5	10	10	1	5

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

Por el contrario, son menos habituales las evaluaciones que abordan el examen de la eficiencia de los programas y proyectos (50%-70%), sus impactos (27%) o la sostenibilidad -impactos duraderos- (27%).

En cuanto a la **Calidad del informe y la utilidad**, las evaluaciones obtienen una elevada puntuación promedio, pues la pauta general es que los informes ofrezcan respuesta adecuada a todas las preguntas (82%)[21].

Asimismo, de forma generalizada presentan de manera balanceada puntos fuertes y débiles de los programas/proyecto (91%; la excepción es PRENATAL); y presentan una distinción clara de declaraciones descriptivas y valorativas (82%)[22], concentrando las segundas en los apartados de conclusiones, por lo general. En el 90% de los casos[23] se especifica como las estimaciones y valoraciones provienen de fuentes confiables, si bien el 27% lo especifica sólo en ocasiones (EDU-INCL, EMPLÉATE y OAPVD). Debe tenerse en cuenta que es habitual que las referencias a la confiabilidad queden establecidas en los apartados de

[21] El 20% obvian alguna pregunta (REDCUDI) o no se aprecia con claridad la respuesta (PRENATAL).

[22] Las excepciones son OAPVD y PRENATAL donde no se aprecian con claridad estas diferencias descriptiva/valorativa.

[23] La excepción es CENCINAI.

metodología, siendo escasas las evaluaciones que acompañan el análisis de datos o la emisión de conclusiones de este tipo de referencias a la confiabilidad desde la que se deben interpretar los resultados.

Tabla 13. Puntuaciones de las evaluaciones en el criterio de Calidad del informe y utilidad

		1 DRAT	2 EDU INCL	3 Germinadora	4 PND	5 REDCUDI	6 CEN-CINAI	7 EMPLEATE	8 OAPVD	9 IMAS	10 PRENATAL	11 PyA-CANCER
4.	**Calidad del informe y utilidad**	***18,8***	***18,8***	***16,5***	***18,8***	***18,8***	***17,8***	***17,5***	***11,8***	***17,8***	***7,5***	***18,8***
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	10	10	10	10	5	10	10	10	10	5	10
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	10	10	10	10	10	10	5	10	10	1	10
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	10	10	10	5	10	10	10	1	10	1	10
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	10	5	10	10	10	1	5	5	10	10	10
4.5	¿Las conclusiones están basadas en los resultados?	10	10	10	10	10	10	10	10	10	1	10
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	10	10	10	10	10	10	10	5	10	1	10
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	10	10	5	10	10	10	10	5	10	1	10
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	5	10	1	10	10	10	10	1	1	10	5

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

Finalmente, en cuanto a calidad de las evaluaciones, cabe destacar buenos registros en la trazabilidad que debe existir entre los hallazgos-conclusiones-recomendaciones, utilizando unos como fundamentos de los siguientes eslabones en la cadena de enjuiciamiento (siendo también fácil recorrer el camino inverso recomendaciones-conclusiones-hallazgos). Esto ocurre en el 82% de las evaluaciones, siendo la excepción PRENATAL, informe en el que no se puede decir que se hayan emitido conclusiones basadas en los hallazgos y en las interpretaciones que sobre ellos formula el equipo evaluador, pues lo que se expone es una síntesis de hallazgos con algún elemento valorativo. También marca la pauta la formulación de recomendaciones relevantes (82%), aunque su grado de especificidad y orientación a la acción (priorización, identificación de destinatarios y pistas para su aplicación) es variable; así como que los informes incorporen resumen ejecutivo (73%)[24].

Y, en cuanto a la evaluación del criterio de **Exhaustividad y adecuación de la documentación** las evaluaciones obtienen una baja calificación, ronzando el incumplimiento su puntuación promedio (5,7).

Tabla 14. Puntuaciones de las evaluaciones en el criterio de Calidad del informe y utilidad

		1 DRAT	2 EDU INCL	3 Germinadora	4 PND	5 REDCUDI	6 CEN-CINAI	7 EMPLEATE	8 OAPVD	9 IMAS	10 PRENATAL	11 PyA-CANCER
5.	**Exhaustividad y adecuación de la documentación**	***6,5***	***5,5***	***6,5***	***5,3***	***5,3***	***4,0***	***6,5***	***4,3***	***5,5***	***5,3***	***8,3***
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	1	1	1	1	1	1	1	1	10	1	0
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	5	1	5	5	5	5	5	1	1	5	5
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	10	10	10	10	5	5	10	10	1	10	10
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos?	10	10	10	5	10	5	10	5	10	5	10

Fuente: elaboración propia a partir de Fichas de registro de evaluaciones.

[24] Las excepciones son Germinadora, OAPVD e IMAS.

En este sentido la pauta de comportamiento queda establecida por una evaluación que no incorpora los TdR en sus anexos (90%)[25]; que incluye sólo algunos de los elementos que especifican el trabajo de campo -cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas- (73%)[26]); y que incluye en anexos la documentación de los instrumentos aplicados (guías para grupos focales, entrevistas guiadas y encuestas) así como una lista de la literatura y documentación utilizada, lo que ocurre en el 91% de los informes, siendo la excepción IMAS.

Finalmente, los informes están estructurados de forma adecuada y su lectura se ve facilitada con gráficos, tablas aunque no suelen incorporar resúmenes de los capítulos (100%), aunque en 4 evaluaciones pueden aportarse mejoras en esta estructura o elementos facilitadores de la lectura[27].

3. Conclusiones

3.1. Valoración del indicador de FOCEVAL

Uno de los objetivos de esta metaevaluación es determinar en qué medida se cumple el indicador de FOCEVAL sobre calidad de las evaluaciones, que se formula como sigue:

"Las evaluaciones realizadas de forma independiente por las contrapartes del proyecto, dentro del marco de medidas de ECD, cumplen con los criterios de calidad internacionales." [TdR de la metaevaluación]

Para emitir un juicio al respecto con base en los hallazgos de esta metaevaluación se presenta la siguiente tabla resumen (Tabla 15) en la que se ubican las evaluaciones en la posición que ocupan respecto al cumplimiento de cada criterio evaluado (alto-medio-bajo). Dado que no existía un estándar de cumplimiento, se han establecido rangos de cumplimiento según la ponderación otorgada a cada criterio en la puntuación final en el esquema de valoración de esta metaevaluación.

Para el criterio de **imparcialidad e independencia** se aprecia que: el 36% de las evaluaciones (4 de 11) se ubica en el tramo de alto cumplimiento, el 45% en el tramo de cumplimiento intermedio (5 de 11), y el 18% (2 de 11) en el tramo de cumplimiento bajo. Cabe apuntar que las evaluaciones que se ubica en el tramo de incumplimiento responden a una evaluación mixta, aunque en su concepción se podría hablar de una evaluación interna-asistida (PRENATAL) y una evalaución participativa (PyA-Cáncer); siendo que el óptimo que implícitamente define este criterio descansa en una evaluación externa. Dado que la configuración del equipo responde a diversos intereses y tiene potencialidades diversas, el reto que se presenta es lograr que todas las evaluaciones en su diseño accedan a una pluralidad de fuentes para mejorar su imparcialidad.

Para el criterio de **equidad, transparencia y participación** se aprecia que: el 55 % de las evaluaciones se ubica en el tramo de alto cumplimiento (6 de 11), el 36% en el tramo de cumplimiento intermedio (4 de 11), y el 9% (1 de 11) en el tramo de cumplimiento bajo o falta de cumplimiento. Cabe apuntar que la única evaluación que se ubica en este tramo inferior responde a una evaluación en la que no hubo participación activa, no hubo taller de inicio con stakeholders o grupos meta y cuyo informe no ha sido accesible a través de internet (IMAS). El reto que vislumbra este resultado es lograr una participación sostenida en todas las evaluaciones para ubicarlas en el tramo superior.

Para el criterio de **credibilidad y exactitud** se aprecia que: el 18 % de las evaluaciones (2 de 11) se ubica en el tramo de alto cumplimiento, y el 82% restante en el tramo de cumplimiento intermedio (9 de 11), destacando el hecho de que no se registra ninguna evaluación en el tramo de incumplimiento o bajo cumplimiento de este criterio clave de evaluación al que se otorga el 50% de la valoración. Siendo esto un buen registro, no es menos cierto que constituye un reto elevar el número de evaluaciones que se ubican en el tramo de alto cumplimiento en cuanto a su credibilidad y exactitud, labor en la que los TdR y los criterios de selección de propuestas juegan un papel relevante, junto a una exigente y cercana labor de gestión de la evaluación.

Para el criterio de **calidad del informe y utilidad** se aprecia que: el 73% de las evaluaciones (8 de 11) se ubica en el tramo de alto cumplimiento, el 9% (1 de 11) en el tramo de cumplimiento intermedio, y el 18% (2 de 11) en el tramo de cumplimiento bajo o incumplimiento. Este criterio se erige en el segundo en importancia

[25] La excepción es IMAS.

[26] El 27% restante no incorpora estos documentos: EDU-INCL, OAPVD e IMAS.

[27] Son los casos de PND, CENCINAI, OAPVD y PRENATAL.

en el esquema de valoración de esta metaevaluación, pues otorga el 20% de la nota final, motivo que permite apreciar la importancia de estos resultados. Cabe apuntar que la única evaluación que se ubica en el tramo de incumplimiento responde a una evaluación (PRENATAL) en la que no se aprecia la exposición de puntos fuertes y débiles de la intervención evaluadas, no se distinguen con claridad declaraciones descriptivas y valorativas y no existe congruencia y trazabilidad en la cadena de enjuciamiento (hallazgos-conclusiones-recomendaciones), y cuyo carácter conclusivo es limitado. El reto en este caso se trata de reforzar la labor de gestión de la evaluación para evitar mermas en los aspectos clave que determinan la calidad y utilidad de un informe de evaluación.

Tabla 15. Grado de cumplimiento de las evaluaciones en los cinco criterios de metaevaluación (alto-medio-bajo)

	IMPARCIALIDAD E INDEPENDENCIA	EQUIDAD, PARTICIPACIÓN Y TRANSPARENCIA	CREDIBILIDAD Y EXACTITUD	CALIDAD DEL INFORME Y UTILIDAD	EXHAUSTIVIDAD Y ADECUACIÓN DE LA DOCUMENTACIÓN	
	10 PUNTOS	10 PUNTOS	50 PUNTOS	20 PUNTOS	10 PUNTOS	
CUMPLIMIENTO ALTO	10 SOBRE 10:	≥ 7,5	≥ 40	≥15	≥ 7,5	
	DRAT Germinadora PND IMAS (Avancemos)	[10.] PyA-CANCER [10] EDU-INCL. [9,0] CENCINAI [8,8] OAPVD [7,8] PRENATAL [7,5] Germinadora	[46,1] DRAT [41,4] IMAS-Avancemos	[18,8] . DRAT [18,8] EDU-INCL. [18,8] PND [18,8] REDCUDI [18,8] PyA-CANCER [17,8] CENCINAI [17,8] IMAS-Avancemos [17,5] EMPLEATE [16,5] Germinadora	[8,3] PyA-CANCER	CUMPLIMIETNO ALTO
CUMPLIMIENTO INTERMEDIO	7,5 SOBRE 10	[5-7,5]	[25–40]	[10–15]	[5-7,5]	
	EDU-INCL. REDCUDI CENCINAI EMPLEATE OAPVD	[6,5] PND [6,0] REDCUDI [5,5] DRAT [5,5] EMPLEATE	[39,3] OAPVD [37,5] Germinadora [34,7] EMPLEATE [33,9] CENCINAI [32,6] . PRENATAL [32,5] PND [31,0] PyA-CANCER [30,0] REDCUDI [27,3] EDU-INCL.	[11,8] OAPVD	[6,5] DRAT [6,5] Germinadora [6,5] EMPLEATE [5,5] EDU-INCL. [5,5] IMAS-Avancemos [5,3] PND [5,3] REDCUDI [5,3] PRENATAL	CUMPLIMIENTO INTERMEDIO
INCUMPLIMIENTO	5 SOBRE 10	< 5	< 25	< 10	< 5	
	PRENATAL PyA-CANCER	[4,6] IMAS-Avancemos		[7,5] PRENATAL	[4,3] OAPVD [4,0] CENCINAI	INCUMPINLIMIENTO

Fuente: Elaboración propia

Para el criterio de **exhaustividad y adecuación de la documentación que incorporan las evaluaciones, ponderado con un 10% de la puntuación,** se aprecia que: ninguna evaluación se ubica en el tramo de alto cumplimiento, concentrándose el 73% en el tramo de cumplimiento intermedio (8 de 11), y el 18% (2) en el tramo de bajo cumplimiento/incumplimiento. Cabe apuntar que las evaluaciones que se ubican en el tramo de incumplimiento (OAPVD y CEN-CINAI) no presentan un cronograma detallado del trabajo de campo realiza o este resulta incompleto, y presentan alguna carencia en cuanto a la estructura del informe o el uso de elemtos que faciliten su lectura (gráficos y tablas, resúmenes, etc.). El reto que vislumbran estos resultados se antoja fácil de alcanzar, incorporando sus requerimientos en los TdR y convirtiéndolos en condiciones para la aprobación de los informes.

Finalmente, si se atiende a la calificación promedio obtenida de la serie de las 11 evaluaciones se aprecia también con claridad el cumplimiento de todos los criterios.

Tabla 16. Grado de cumplimiento de las evaluaciones en los cinco criterios de metaevaluación (alto-medio-bajo)

Criterios	Subcriterios /Dimensiones	Ponderación	PROMEDIO
1. Imparcialidad e independencia:		*10 puntos*	**8,0**
2. Equidad, participación y transparencia:		*10 puntos*	**7,4**
3. Credibilidad y exactitud:		*50 puntos*	**35,1**
3.1	Metas, objetos y modelo de cambio	[*10p.*]	**6,9**
3.2	Diseño y métodos	[*10 p.*]	**7,8**
3.3	Procedimiento de selección	[*10p.*]	**6,6**
3.4	Análisis de datos	[*10p.*]	**6,2**
3.5	Criterios de evaluación	[*10p.*]	**7,6**
4. Calidad del informe y utilidad:		*20 puntos*	**16,6**
5. Exhaustividad y adecuación de la documentación		[*10p.*]	**5,7**
		TOTAL	71,3

Fuente: Elaboración propia

Se puede afirmar por tanto que las evaluaciones cumplen el indicador fijado por FOCEVAL tomando como estándar de cumplimiento la puntuación promedio de 71,3 sobre los 100 puntos totales del esquema de valoración de los criterios y preguntas de metaevaluación. Esto se ve reforzado por varios hechos:

1. En el criterio de credibilidad y exactitud (el más relevante, 50%) no se ubica ninguna evaluación en el tramo de incumplimiento/cumplimiento bajo.

2. En el criterio de calidad y utilidad del informe (segundo en relevancia, 20%), tan sólo una evaluación se ubica en el tramo de cumplimiento bajo/incumplimiento.

3. De las 4 evaluaciones que se ubican en el tramo inferior de alguno de los criterios, sólo una acumula este registro en dos criterios, apuntando a que en la mayoría se trata de una debilidad "razonable".

4. El criterio que registra menos evaluaciones en el tramo superior puede considerarse un criterio que si bien supone una merma de calidad del informe para que sus lectores puedan juzgar adecuadamente la solvencia de sus conclusiones y recomendaciones, eso no significa que los gestores no dispongan de dicha información y pueda ser, de alguna manera, subsanable.

3.2. Factores determinantes de la calidad general de las evaluaciones.

Los TdR establecían que el capítulo de apreciación genreal de la calidad de las evaluaciones integrase la calificación de los factores centrales que determinan la calidad de las evaluaciones y los procesos evaluativos. Cabe apuntar que resulta complejo identificar estos factores, por un lado debido a la diversidad de contextos evaluativos y modalidades de apoyo bajo la que se desenvolvieron las evaluaciones. Por otro lado, el hecho de contar sólo con el análisis documental, plantea limitaciones para interpretar adecuadamente aquellos factores que incidieron en mayor medida en la calidad de los procesos.

No obstante, el análisis de los informes de evaluación (Fichas de registro y Síntesis) permite apreciar dificultades compartidas los equipos de evaluación (calidad de la información disponible o exigencias de las modalidades de apoyo). En otros casos, estos factores no se muestran directamente, sino que son fruto de la interpretación que sobre los datos recabados formula el metaevaluador (focalización de los TdR, disponibilidad de recursos o la participación en los procesos).

Por estos motivos, en este apartado se exponen los factores que a la luz de los criterios de metaevaluación destacan como determinantes de la calidad de las evaluaciones; añadiendo "otros factotres" que a juicio del metaevaluador se interpreta que han tenido incidencia en la calidad de los procesos.

Factores objetivos determinantes de la calidad según los criterios de metaevaluación.

- **Imparcialidad e independencia.**
  - **La bondad de la implementación de los mecanismos de apoyo al proceso evaluativo.** No cabe duda de que aquellas evaluaciones que cuentan con mecanismos de apoyo específicos para la mejora de su calidad encuentran en estos (o deberían encontrar) un factor determinante de su calidad. El

asesoramiento a los equipos de gestión de las evaluaciones o a los equipos de evaluación, o la conformación de equipos de evaluación expertos o académicos, constituyen modalidades de apoyo experimentadas que debieran erigirse en factores clave. Si bien, es difícil apreciar este vínculo, hasta tres equipos de evaluación[28] han mostrado las dificultades que entraña implementar la modalidad de aprender-haciendo, pues las cargas de trabajo de los gestores y el tiempo que requieren estos procesos, se han apuntado como factores que han afectado a los procesos, especialmente en cuanto a la optimización de la gestión de los plazos disponibles para realizarlas con un adecuado desarrollo de capacidades, así como por el sobre esfuerzo que requiere para quienes prestan las asesorías, escasamente reconocido en los presupuestos, según se apuntó, lo que acaba traduciéndose en mermas en la calidad y en que los propios gestores sientan la necesidad de haber requerido un mayor asesoramiento del recibido. Estos procesos también han generado interferencias, según se apuntó en algún caso, con la independencia e imparcialidad de la evaluación, relacionada también con la participación en el proceso de las instituciones responsables de los programas evaluados, lo cual exige una cuidadosa delimitación de roles y funciones a lo largo del proceso para evitarlas.

- **Equidad, participación y transparencia:**
  - **El grado de participación efectivo y sostenido en el proceso evaluativo**. La participación constituye un determinante de primera magnitud de la calidad de los procesos evaluativos, entendida esta como potencial que poseen sus recomendaciones para ser aplicadas por sus destinatarios, pues en última instancia condiciona su utilidad real y su uso efectivo.
    - ¿*Un trade-off participación-credibilidad*?. En esta metaevaluación hubo evaluaciones que obtuvieron una alta valoración en términos de credibilidad y exactitud o calidad del informe, siendo paradigmático el caso de DRAT, pero para las que no se registró participación, siendo que se presentan dudas razonables sobre el uso efectivo que se haya podido efectuar de sus resultados. Aunque no puede aventurarse una correlación entre calidad o credibilidad y participación a partir de los hallazgos de esta metaevaluación[29], aquellos procesos que registran una baja o nula participación en su diseño o en su implementación se presentan como evaluaciones en las que la aplicación de las recomendaciones encontrará un obstáculo importante, pues es probable que no se haya producido un ajuste en las expectativas e intereses de los actores clave, ni estos se hayan ido apropiando del proceso como antesala para asumir las recomendaciones, tanto más, cuánto más sustantivas sean en los cambios que requiere la intervención evaluada para su mejora (cambios en la propia conceptualización del programa), espacios donde suelen terminar aplicándose recomendaciones de escasa relevancia, de naturaleza operativa o procedimental. Al contrario, emerge el hecho de que evaluaciones participativas genuinas generen un buen caldo de cultivo para inducir mejoras, pero que sus limitaciones en credibilidad y exactitud le resten argumentos.
- **Credibilidad y exactitud.**
  - **El uso efectivo de las teorías de la intervención diseñadas y la explicitación de las hipótesis de impacto.** Un factor que contribuye a la calidad de la evaluación es la explicitación de la teoría en la que se fundamenta la intervención, pues constituye uno de los principales focos del aprendizaje para la mejora. Dado que estas teorías no suelen estar explícitas es relevante que se (re)construyan durante los procesos evaluativos. En las evaluaciones metaevaluadas se ha erigido como un factor positivo el hecho de que se hayan (re)construido estas teorías, aunque esto era algo que expresamente

[28] EDU-INCL, Germinadora y EMPLEATE.

[29] No obstante, resulta relevante observar en la anterior tabla la posición que ocupan determinados proyectos en el criterio equidad, participación y transparencia (DRAT o IMAS) y la que ocupan en el criterio de Credibilidad y exactitud. Esto permite aventurar una hipótesis que debiera contrastarse en próximos ejercicios acerca de que las evaluaciones con menos participación pueden prestar mayor atención a los propios requerimientos que exige el rigor técnico, si bien es deseable que una evaluación obtenga una posición balanceada de ambos, pues cabe esperar que de lo contrario el volumen de recomendaciones estratégicas aceptadas y aplicadas sea relativamente bajo. En el extremo opuesto se encuentran PyA-Cáncer y EDU-INCL, dos evaluaciones que registraron la máxima puntuación en participación, equidad y transparencia, mientras que se ubican en las últimas posiciones en términos de credibilidad y exactitud. En este caso, si bien la participación alcanza hace presagiar una buena utilidad/utilización de la evaluación, se debe prestar mayor atención a su credibilidad y exactitud.

solicitaban los TdR, que prestaron especial atención a este aspecto. No obstante, el hecho de que no se hayan aplicado en general como ejes vertebradores de los procesos ha limitado sus registros en calidad, pues no han desplegado todo el pontencial que tienen y no se ha facilitado a gestores y planificadores su contrastación como parte del ejercicio evaluativo, ni se han dirigido específicamente recomendaciones de manera sigfinitiva orientadas precisamente a mejorar o corregir dichas teorías a partir de las evidencias recabadas.

- **La atención prestada al diseño metodológico de las evaluaciones.** Un factor determinante de la calidad de las evaluaciones es el que hace referencia a la solvencia de sus fundamentos metodológicos. Las evaluaciones metaevaluadas han prestado desigual atención a la metodología, entendida esta como el estudio de los métodos o teorías en las que se sustenta la evaluación y la reflexión sobre la forma idónea para generar evidencias en la evalaución y sobre el porqué de la secuencia en la que se aplicarán las diferentes técnicas en el proceso evaluativo. Igualmente, se prestó escasa atención en general –con excepciones- a los enfoques y modelos teóricos de evaluación, como recursos muy relevantes para orientar la evaluación hacia sus intereses explicitados. En ausencia de un mayor desarrollo metodológico, aquellas evaluaciones que han mostrado mayor solvencia y sistematicidad en el enjuiciamiento han sido aquellas que presentan una matriz de evaluación y ésta ha sido efectivamente utilizada para vertebrar tanto los hallazgos como sobre todo las conclusiones y recomendaciones (criterios y preguntas). En el extremos opuesto se situan aquellas que o bien no presentan dicha matriz o en las que esta ha vertebrado en desigual forma sus conclusiones y recomendaciones. Aunque no puede circunscribirse un diseño metodológico a una matriz de evaluación, cabe destacar la importancia que estas adquirieron como determinantes de la calidad de las evaluaciones. Si bien los TdR pueden ofrecer alguna pista en este sentido (como bien hacen los de CENCINAI), y que los equipos gestores de las evaluaciones pueden tratar de que se fortalezcan estos contenidos y reflexión metodológica en los procesos evaluativos, lo cierto es que es en última instancia la pericia de los equipos de evaluación la que determinará un uso y manejo óptimo de estos enfoques y modelos.

- **La disponibilidad y calidad de la información de monitoreo**. Este factor determina no sólo la posibilidad de obtener hallazgos suficientes y solventes cuando uno de los focos evaluativos son los procesos de implementación o ejecución, sino que se ha puesto de relieve que facilita la labor de elección de las muestras para la encuestación, haciendo mucho más eficiente la evaluación. Asimismo, se han revelado como un factor clave a la hora de evaluar la relevancia y pertinencia, pues son la base para caracterizar a la población beneficiaria y sus necesidades; así como para evaluar la eficacia y poder emitir juicios concluyentes a cerca de los impactos y la sostenibilidad de los mismos; permitiendo la contrastación de las teorías de la intervención diseñadas y de las hipótesis de impacto establecidas.

- **La idoneidad de las muestras utilizadas y de los procedimientos de selección.** Cuando se recurre a encuestas en los procesos evaluativos, no cabe duda de que la bondad de las muestras que se contemplen se erige en un factor determinante de su calidad, tanto más cuanto más protagonismo se le otorgue a esta técnica en la matriz de evaluación, pues determinan la fiabilidad de los datos y hallazgos que sustentan las conclusiones y recomendaciones. No cabe duda también, que la idoneidad de las encuestas realizadas (muestras y errores) es un factor muy asociado a los recursos disponibles y a la cantidad y calidad de la información disponible (bases de datos, monitoreo, etc.). En este sentido se aprecia en general en los informes que las evaluaciones aceptaron errores muestrales elevados, en cuya base está a buen seguro una decisión de factibilidad dados los recursos e información disponibles (como causas asociadas a este factor).

Calidad del informe y utilidad.

- **La congruencia y buena trazabilidad entre hallazgos, conclusiones y recomendaciones.** Es este un factor determinante de la calidad de las evaluaciones en tanto que genera un procedimiento de enuciamiento sistemático, en el que no tienen cabida los prejuicios de los equipos de evaluación, sino que toda conclusión se apoya en un hallazgo del proceso, y toda recomendación se ancla en una recomendación. Ha sido este un factor positivo de las evaluaciones que ha contribuido en gran medida a obtener buenos registros en este criterio, pues la mitad de las preguntas que lo concretan se refieren a esta cuestión (4.4 a 4.7).

- ⇨ **La respuesta adecuada a las preguntas de evaluación.** Es este un factor clave como determinante de la calidad de las evaluaciones, pues si puede decirse que estas comienzan con alguien que se hace preguntas sobre una intervención pública, es la respuesta que encuentran la que determina su calidad. En este sentido, ha sido este un factor que ha actuado positivamente en las evaluaciones objeto de metaevaluación y que ha contribuido a sus buenos registros en términos de calidad y utilidad (potencial).

- **Exhaustividad y adecuación de la documentación.**
  - ⇨ **La atención prestada a la documentación que anexan las evaluaciones.** Es este también un factor que contribuye a la calidad de las evaluaciones, sobre todo, en tanto que fortalece su solvencia, facilita la trazabilidad de sus hallazgos y conclusiones, y contribuye a aligerar el contenido de los informes haciéndolos más manejables y útiles mediante el uso de referencias a sus anexos en los que se puede encontrar mayor desarrollo de aspectos técnicos. En este sentido, ha sido un factor que en general ha restado calidad a las evaluaciones metaevaluadas, pues prestaron una atención muy desigual a estos contenidos. Dicho esto, es precisamente el criterio en el que la mejora y la subsanación de deficiencias se antoja más sencilla, pues tan sólo el hecho de que se establezca y explicite de antemano el contenido de deseado (en los TdR), puede ser fácilmente contemplado por los equipos de evaluación y supervisado por los propios gestores de los procesos evaluativos.

Otros factores.

- **Los Términos de referencia: su focalización y sus requerimientos**. Se observa que muchos de los aspectos de mejora de las evaluaciones se corresponden con ámbitos que no encontraban especificación o énfasis en los TdR (documentación anexa, resumen ejecutivo, estructura del informe, criterios y procedimientos de selección, taller de inicio, etc.). Asimismo, la focalización del encargo en pocas preguntas relevantes ayuda sin duda a generar un encargo evaluativo mucho más concreto.
- **Los procesos de gestión de la evaluación**. La supervisión de la calidad de los procesos que realizan los equipos gestores de las evaluaciones es también un determinante de la calidad de la evaluación, en tanto que deben ir exigiendo aquellos parámetros que determinan la calidad del proceso, sobre todo en aquellos casos en los que se trata de aspectos que difícilmente pueden quedar especificados en unos TdR (errores muestrales, fomento de la participación, integración de la teoría de la intervención en la evaluación, diseño metodológico, o exhaustividad de la documentación de los informes, entre otros).

4. Recomendaciones derivadas de la metaevaluación

En este apartado se presentan las recomendaciones que se derivan de esta metaevaluación. Su finalidad es producir resultados que contribuyan a la mejora de los procesos de evaluación en Costa Rica (de MIDEPLAN y otros actores); así como servir de insumo para dos talleres que según especifican los TdR servirán a FOCEVAL para elaborar las lecciones aprendidas sobre cómo mejorar la calidad de las evaluaciones, el proceso evaluativo, el uso de las evaluaciones y la interacción entre estos puntos.

Las recomendaciones presentan tres focos de atención como especificaban los TdR: mejora de la calidad de las evaluaciones y TdR, los procesos evaluativos y –si posible–utilidad de las evaluaciones.

4.1. Calidad y utilidad de las evaluaciones y los procesos evaluativos

Estas recomendaciones, que se presentan priorizadas según su importancia, se destinan al proyecto FOCEVAL en general y a MIDEPLAN en particular, como institucionalidad púlbica que ejerce la rectoría del Sistema Nacional de Evaluaciones.

1. **Desplegar medidas para optimizar la calidad del diseño metodológico de las evaluaciones**, promoviendo la reflexión de los equipos de evaluación e incentivando la creatividad en el uso de enfoques y modelos[30]. Aunque no hay una pregunta expresa sobre este aspecto (implícitamente en la idoneidad

[30] La conclusión que alimenta esta recomendación, puede considerarse que es un parámetro recurrente en las metaevaluaciones. Al menos en las que ha realizado el autor: México DF, Portugal y España. Al ser concebidos como trabajos de consultoría y responder a

del ajuste del diseño de la evaluación a los intereses de la misma), se aprecia de manera generalizada que las evaluaciones prestan poca atención a los fundamentos metodológicos y a los enfoques y modelos teóricos como ejes de los procesos evaluativos. Unas pasan directamente de las preguntas a la exposición de la matriz de evaluación o al diseño de técnicas de investigación y otras enuncia algún referente teórico que no se aprecia en su implementación. Se recomienda:

a) Invertir en capacitación especializada y selectiva de la comunidad evaluadora, prestando especial atención a los enfoques teóricos y sus implicaciones en los procesos evaluativos (la evaluación basada en la teoría se presenta adecuada para que las teorías de la intervención que se vislumbran en los informes, se erijan en ejes vertebradores de los procesos).

b) Prestar atención a este aspecto en los TdR**,** como una manera de inducir su uso en las propuestas y en los procesos e informes de evaluación.

2. **Desplegar medidas para optimizar la participación efectiva y de calidad en los procesos de evaluación y en todas sus etapas.** La participación está íntimamente ligada a la utilidad de los procesos de evaluación, tanto por el aprendizaje sobre la marcha (aprender-haciendo), como para despertar el interés de los destinatarios de los informes, y, en última instancia para facilitar la adopción de recomendaciones de mejora, tanto más cuanto más sustantivas o estratégicas sean estas. Entre estas medidas cabe especificar algunas como:

   a) Prestar atención específica en los TdR, estableciendo indicaciones precisas sobre el tipo de actores a implicar, pero también sobre los momentos clave (por ejemplo, taller de inicio)

   b) Desplegar procesos participativos para la propia elaboración de los TdR (y reflejarlos en éstos –por ejemplo la necesidad de realizar un taller de inicio para la sensibilización-),

   c) Generar espacios de diálogo entre los equipos de evaluación y los principales agentes implicados en estructuras formales de participación para facilitar la reflexión colectiva para la mejora (y preverlo en los TdR, al menos al inicio, a medio camino y a la entrega del Informe borrador).

3. **Desplegar medidas para impulsar la mejora de la evaluabilidad de los programas y proyectos para los que se propone su evaluación.** Entre estas medidas, cabe especificar una en el marco del proceso de selección de programas que conformarán la ANE, pues aunque en los protocolos establecidos desde MIDEPLAN se contempla la evaluabilidad, los problemas que han reflejado en sus informes los equipos de evaluación aconsejan tanto su mejora, como la necesidad de dedicar mayores esfuerzos a impulsar un proceso generalizado de la mejora de evaluabilidad de los programas y proyectos, con énfasis en la necesidad de disponer de mejores sistemas de información orientados a la evaluación, pues en aquellos casos en los que estos existen, presentan muchas deficiencias para la evaluación, consumiendo grandes recursos para los equipo de evaluación (depuración de bases de datos, diseño de muestras, etc.). Para optimizar estos esfuerzos sería recomendable que los propios TdR realicen algunas especificaciones sobre las expectativas o el tipo de recomendaciones que se esperan, entre las que resultaría relevante incluir aquellas relativas a la mejora de la evaluabilidad (contribuiría a visibilizar esta problemática e incorporarla sistemáticamente en los Planes de Acción que se derivan de las evaluaciones).

4. **Impulsar el uso de las teorías de la intervención como ejes vertebradores de los procesos evaluativos**, de manera que se evite la inversión que supone su construcción colectiva para que erija en un subproducto de los procesos. Donde reside su potencial para el aprendizaje organizacional y la mejora de las políticas públicas es precisamente en su contrastación mediante los ejercicios evaluativos.

5. **Establecer un soporte documental -mínimo y deseado- para los procesos de evaluación,** de manera que se induzca un salto de calidad en la exhaustividad y adecuación de la información y documentación que anexan los informes (una debilidad compartida por los informes metaevaluados) y, a la vez, se ofrezca una pista suficiente para la actividad metaevaluadora (evitando disponer sólo de un 30% de los TdR como en esta metaevaluación). Los expedientes de los procesos evaluativos licitados deben documentarse con los TdR, las propuestas técnicas/ofertas presentadas, los informes de evaluación y sus anexos, y las respuestas de gestión emitidas por los equipos de gestión de la evaluación sobre la aceptación de los

unos requerimientos de técnicas de unos TdR, los equipos de evaluación suelen prestar poca atención a estos aspectos, aun cuando implícitamente se guíen por un determinado enfoque o modelo.

diferentes productos contemplados y sobre la valoración de la calidad de las propuestas técnicas presentadas y del informe final (matrices de puntajes).

6. **Identificar buenas prácticas o puntos estrella en las evaluaciones realizadas** de manera que se erijan en referente para los evaluadores, y para los gestores. La matriz de puntajes de esta metaevaluación puede ayudar en ese proceso de identificarlos.

7. **Optimizar las modalidades de apoyo a las evaluaciones, especialmente en el caso de Aprender-haciendo.** Se hace necesario contemplar un mayor horizonte temporal en estos procesos para que se produzca un adecuado desarrollo de capacidades, así como una cuidadosa delimitación de roles y funciones entre los integrantes del equipo gestor-evaluador, sobre todo atendiendo a sus instituciones de procedencia y su vinculo directo con el programa o proyecto evaluado.

8. **Impulsar la concurrencia competitiva en los procesos de evaluación** de manera que se promueva la mejora de calidad de las evaluaciones. Los TdR analizados así como las referencias al respecto en los informes de evaluación, ponen de relieve la importancia que adquirieron los procesos de contratación directa o en los que la concurrencia estuvo limitada. Aunque no cabe apuntar a priori que las estas evaluaciones así seleccionadas tengan menor calidad, y esto pueda estar más justificado en algún caso (evaluaciones académicas) se espera que la mayor concurrencia estimule la competitividad y eleve la calidad de las evaluaciones. Esta recomendación lleva aparajeda la mayor difusión posible de las licitaciones de las evaluaciones de manera que se permita atraer mayor número de licitadoes.

9. **Concentrar los esfuerzos de los evaluadores en la producción de recomendaciones relevantes y específicas para los actores,** más que dispersarlos en la producción de planes de acción y socialización, tareas que si aquellas son de calidad, pueden asumir los gestores con mayor conocimiento y efectividad. Son pocas las evaluaciones que finalmente los contemplan, aunque los hay muy meritorios como en la evaluación del PND, que puede conformar un banco de bunas prácticas.

10. **Clarificar la composición de los equipos implicados en los procesos evaluativos.** Para evitar equívocos, se recomienda que los informes de evaluación clarifiquen en su portada las entidades y personas que participaron en la evaluación diferenciando con claridad al menos tres roles: equipo evaluador (Coordinador e integrantes), equipo gestor de la evaluación, y enlace institucional[31].

4.2. Calidad de los Términos de Referencia

Estas recomendaciones se orientan especialmente a MIDEPLAN y en el marco de la ANE.

11. **Homogeneizar la estructura y visión de los TdR de las evaluaciones en el marco de la Agenda Nacional de Evaluaciones donde el rol director de MIDEPLAN y de aquellas otras que se impulsen desde MIDEPLAN y en el marco del proyecto FOCEVAL,** tomando como referente los publicados para CEN-CINAI, pues resultan los más completos, claros y ordenados, de los analizados. Para hacer extensiva esta práctica, resulta relevante disponer de un modelo de TdR en el marco de la ANE, que sea accesible para los gestores de los programas y para los gestores de evaluaciones.

12. **Incorporar y desarrollar orientaciones sobre el diseño metodológico de las evaluaciones en los TdR.** Esta recomendación no trata de impulsar evaluciones en las que su metodolgía esté "prefijada", muy al contrario lo que persigue es favorecer la libertad y creatividad de los equipos de evaluación en este sentido (como bien se hace en los TdR de CEN-CINAI), reforzando y concretando aquellos aspectos del diseño metodológico que se espera que desarrolle el equipo evaluador en su propuesta técnica y en la evaluación de ser adjudicatario (reflexión sobre los métodos, teorías en las que se sustenta, enfoques o modelos teóricos de evaluación por los que se opta o diferenciando el diseño metodológico en sí de la construcción de matrices de evaluación).

13. **Incorporar en los TdR de manera detallada la documentación que se espera que incorporen los informes finales,** tomando como referente aquella que especifican los criterios de esta metaevaluación: anexar los TdR, guias de entrevistas y grupos focales, cuestionarios, docuemntación utilizada, cronograma detallado del trabajo de campo (lugares, personas y fechas), así como sus propios TdR.

[31] Ver Evaluación 02 DRAT, que en su página 2 hace una aclaración acertada en esta línea.

14. **Establecer orientaciones sobre las expectativas de recomendaciones que se esperan en los TdR,** siendo un buen ejemplo CENCINAI. Entre las indicaciones que deben establecerse cabe apuntar la necesidad de establecer priorizaciones en las recomendaciones que se formulan, identificar la entidad u órgano destinatario principal, así como ofrecer pistas que faciliten su aplicación a los gestores.
15. **Contemplar en los TdR el acceso a la consulta de los criterios a partir de los que se juzgará la calidad de las evaluaciones** apoyadas de manera que se permita a los equipos de evaluación su consulta.
16. **Evitar encargos de evaluación que generen la necesidad de muchas entregas o subproductos en los procesos de evaluación, de manera que el equipo evaluador pueda concentrarse en el Informe final[32].**

4.3. Calidad y utilidad de la metaevaluación

Estas recomendaciones se orientan a FOCEVAL y tienen como propósito contribuir a optimizar el uso de la metaevaluación así como la posibilidad de realizar nuevos ejercicios similares.

17. **Revisar y ajustar los criterios y preguntas, tanto su formulación precisa y su estructura** para evitar duplicidades y facilitar la correcta interpretación de aquello que se desea saber con cada uno de ellos. Asimismo, es importante otorgar más peso y visibilidad en los criterios y preguntas a la calidad de los TdR, pues actualmente se juzgan de manera indirecta y difusa.
    a) Contemplar en el criterio de Imparcialidad e independencia la concurrencia competitiva del proceso de contratación de la evaluación, otorgando mayor valor a los que si la tuvieron.
18. **Contemplar una mayor dimensión de los ejercicios de metaevaluación[33]**
    a) Incorporando trabajo de campo, para poder apreciar con mayor fundamento los procesos evaluativos, más allá de sus productos (Informes, anexos, …). En este sentido, para poder hablar en puridad de que se está juzgando la calidad del proceso y no sólo del producto, debe incorporarse la perspectiva de los gestores de evaluaciones pues, son en última instancia, en quienes descansan los procesos de ECD y quienes interactúan con los equipos de evaluación para optimizar la calidad.
    b) Incorporando el trabajo en equipo más allá de la consultoría individual para permitir el contraste de juicios (peer review,…), minimizando sesgos de una única lectura e interpretación del proceso.
    c) Incorporando el seguimiento a las recomendaciones, de manera que se aporten mayores evidencias sobre el uso efectivo de las evaluaciones, generando así hallazgos más solventes para obtener conclusiones más certeras sobre la utilidad.
19. **Publicar los criterios, dimensiones y preguntas de metaevaluación y referenciarlas en los TdR de las evaluaciones,** de forma que tanto gestores de las evaluaciones como evaluadores conozcan de antemano los parámetros a partir de los cuales se juzgará el valor y el mérito de sus evaluaciones.

[32] Un ejemplo ilustrativo de TdR que solicita excesivas entregas intermedias, que consumen mucho tiempo del equipo evaluador en su elaboración, edición y maquetación, etc. es el caso de REDCUDI.

[33] Es cierto que desde FOCEVAL se han puesto en marcha otras actividades y consultorías orientadas a recabar este tipo de percepciones cualitativas sobre los procesos evaluativos y para la sistematización de experiencias, pero no cabe duda que desde el punto de la metaevaluación en sí, le resta potencial como instrumento para impulsar la mejora de la calidad de los procesos evaluativos, el hecho de que no se pueda completar su indagación con el trabajo de campo (entre ellos resultan especialmente pertinentes las entrevistas y grupos focales con gestores y evaluadores, encuestas a gestores y evaluadores).

5. Anexos

5.1 Términos de referencia de la metaevaluación

Términos de Referencia

Realización de una metaevaluación

«Experto en evaluación para realización de una metaevaluación»

(Redacción de 22.08.2016)

1. Contenido

1a) Antecedentes

El Instituto Alemán de Evaluación de la Cooperación al Desarrollo (Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit, DEval) se fundó por iniciativa del Ministerio Federal de Cooperación Económica y Desarrollo (Ministerium für wirtschaftliche Zusammearbeit und Entwicklung, BMZ) y emprendió su trabajo en el otoño de 2012. El Instituto tiene la forma legal de una sociedad de responsabilidad limitada sin ánimo de lucro y está inscrito en el Registro Mercantil de Bonn. Como socio único figura el Estado Federal, representado a través del Ministerio Federal de Cooperación Económica y Desarrollo (Ministerium für wirtschaftliche Zusammearbeit und Entwicklung, BMZ).

La tarea principal del DEval consiste en el análisis y la valoración (evaluación) independiente de la actividad pública alemana de cooperación al desarrollo. Otras de sus tareas incluyen el desarrollo de métodos y estándares adecuados, así como de medidas de cualificación, y la potenciación de los recursos de evaluación (evaluation capacity development, ECD) en los países en los que se desarrolla la actividad alemana de cooperación al desarrollo. A tal fin, el Instituto coopera a escala nacional e internacional.

1b) El proyecto FOCEVAL

FOCEVAL (Fomento de Capacidades en Evaluación en Costa Rica y países seleccionados de América Latina) es un proyecto de fortalecimiento de capacidades en evaluación financiado con fondos del BMZ implementado por DEval, la contraparte principal del proyecto es el Ministerio de Planificación Nacional y Política Económica de Costa Rica (MIDEPLAN).

El proyecto FOCEVAL fue concebido como una plataforma interinstitucional para apoyar la creación de capacidades en evaluación tanto en Costa Rica como en otros países de América Latina. Se gestiona FOCEVAL desde Alemania, pero tiene un equipo de coordinación local, asistido técnicamente por los evaluadores y personal de DEval que opera desde Bonn (Alemania) y que realiza visitas periódicas a Costa Rica.

El proyecto tiene el objetivo de fortalecer el rol que juega la evaluación en Costa Rica y en países seleccionados de América Latina en la toma de decisiones y en la conducción política.

Para lograr este objetivo, el proyecto tiene tres componentes: capacitación, regionalización e institucionalización. En este último se enmarca como una actividad principal el acompañamiento de la ejecución de la Agenda Nacional de Evaluación en Costa Rica.

1c) La Agenda Nacional de Evaluaciones del Plan Nacional de Desarrollo 2015-2018 de Costa Rica

Siguiendo la recomendación surgida de la evaluación del diseño y proceso de elaboración del Plan Nacional de Desarrollo (PND) 2011-2014, el Plan Nacional de Desarrollo 2015 - 2018 incluye una Agenda Nacional de Evaluaciones (ANE). La ANE agrupa las intervenciones públicas contenidas en el PND que los sectores, en conjunto con las instituciones que los componen, identificaron como estratégicas y en virtud de ello se proponen para ser objeto de evaluación.

El órgano responsable del diseño e implementación de la ANE es MIDEPLAN, institución rectora del Sistema Nacional de Evaluación de Costa Rica.

FOCEVAL apoya la implementación de varias evaluaciones de la ANE. Para aprender cómo mejorar la calidad de las evaluaciones, el proceso evaluativo y el uso de las evaluaciones (y la interacción entre estas dimensiones), se quiere celebrar dos talleres a principios del 2017. Como un insumo importante para estos talleres se requiere hacer una metaevaluación de por lo menos 10 evaluaciones realizadas en Costa Rica en los últimos años.

2. Objetivos y tareas del encargo

DEval pretende contratar a un experto que cuente con varios años de experiencia en el área de la planificación, la gestión y la implementación de evaluaciones para realizar una metaevaluación de las evaluaciones elaboradas dentro de la ANE y otras evaluaciones de programas nacionales en Costa Rica, incluyendo los respectivos términos de referencia.

La metaevaluación debe cumplir con tres objetivos:

a) Valorar el indicador de FOCEVAL “Las evaluaciones realizadas de forma independiente por las contrapartes del proyecto, dentro del marco de medidas de ECD, cumplen con los criterios de calidad internacionales.”

b) Producir resultados que contribuyan a la mejora de los procesos de evaluación en Costa Rica (de MIDEPLAN y otros actores).

c) Elaborar resultados y recomendaciones que sirvan como insumo para dos talleres en los que se elabora lecciones aprendidas sobre cómo mejorar la calidad de las evaluaciones, el proceso evaluativo, el uso de las evaluaciones y la interacción entre estos puntos.

Para ello, la persona experta contratada deberá valorar la calidad metodológica, la validez de los resultados y los términos de referencia[34] de un mínimo de 10 evaluaciones según 30 a 40 criterios pre-establecidos, que corresponden a estándares internacionales y que serán compartidos con la/el experta/o después de haber firmado el contrato.

[34] Hay evaluaciones sin términos de referencia, entonces no se valorizará los términos de referencia.

2a) Las evaluaciones evaluados en la metaevaluación

La metaevaluación incluirá principalmente las siguientes evaluaciones:

1. Evaluación de Impacto del "Distrito de Riego Arenal-Tempisque", de Servicio Nacional de Áreas Subterráneas, Riego y Avenamiento (SENARA).

2. Evaluación del proyecto "Apoyo a la Gestión pedagógica a Centros educativos de calidad con orientación inclusiva", del Ministerio de Educación Pública (MEP).

3. Evaluación del proyecto interinstitucional "Germinadora de Desarrollo Organizacional Empresarial, Asociativo y Comunitario".

4. Evaluación del "Diseño Metodológico y Proceso de Elaboración del PND 2011 – 2014", de MIDEPLAN.

Estas cuatro evaluaciones (piloto) fueron realizadas en la primera fase de FOCEVAL por consultores nacionales e internacionales.

5. Evaluación del programa interinstitucional "Red Nacional de Cuido y Desarrollo Infantil (REDCUDI)".

6. Evaluación del programa "Centros de Educación y Nutrición, y Centros Infantiles de Atención Integral (CEN-CINAI)", del Ministerio de Salud.

7. Evaluación del programa "Empléate", del Ministerio de Trabajo y Seguridad Social.

Estas tres evaluaciones fueron realizadas en la segunda fase del proyecto FOCEVAL en el marco de la ANE. Las evaluaciones de los programas CEN-CINAI y REDCUDI fueron ejecutadas por evaluadores nacionales y gestionadas por un "equipo gestor" compuesto por funcionarios y funcionarias de las instituciones involucradas en los programas. Los equipos gestores fueron asesorados por consultores internacionales. La evaluación del programa Empléate fue realizada bajo la modalidad aprender-haciendo, en el que el mismo equipo gestor fue el equipo evaluador y realizó la evaluación con el acompañamiento de dos consultoras.

8. Evaluación del "Programa de Atención de la Oficina de Atención y Protección a la Víctima de Delito", del Poder Judicial de Costa Rica.

9. Evaluación del programa "Avancemos" del Ministerio de Educación Pública, junto con otras instituciones

Estas fueron dos evaluaciones académicas de aprendizaje ejecutadas por la MEPPD (Maestría en Evaluación de la UCR) jóvenes y emergentes evaluadoras, con el apoyo de FOCEVAL durante su primera etapa (en el primer caso) y el BID (en el segundo caso).

10. Ministerio de Salud: a ser definida.

Además de estas 10 evaluaciones, en este momento FOCEVAL está en el proceso de identificar otras evaluaciones concluidas en los últimos años. En este momento el proceso no se cerró todavía. En el caso que se incluirán otras evaluaciones más se considera una ampliación del contrato, aumentando los días previstos.

3. Descripción de la prestación a realizar

El/La **contratista** realizará las siguientes prestaciones:

El contratista valora las evaluaciones mencionadas según criterios preestablecidos que serán compartidos una vez firmado el contrato. La metaevaluación debe valorar la metodología aplicada en las evaluaciones así como la validez de sus resultados. En la metaevaluación el contratista considerará los informes finales de las evaluaciones y los Términos de Referencia (TdR) de las evaluaciones (cuando corresponda), ya que éstos influencian, en buena medida, la calidad de la evaluación. Además, la metaevaluación deberá determinar en qué medida la evaluación correspondió a lo demandado en los TdR.

La metaevaluación deberá ser presentada en un informe final en el que se incluye una calificación general de las evaluaciones y se describe y valora la calidad de cada evaluación y sus TdR según los criterios preestablecidos. En el informe también se deberá realizar recomendaciones referentes a cómo mejorar la calidad de las evaluaciones y el proceso evaluativo, incluyendo la interacción entre estos puntos.

En ese sentido, las recomendaciones de la metaevaluación deberán ser concretas y enfocadas para poder contribuir a la mejora de procesos de evaluación de

MIDEPLAN. Además la metaevaluación debe formular una calificación clara de las evaluaciones, para medir el indicador mencionado bajo el punto 2), objetivo a) (ver arriba).

El informe final de la metaevaluación debe tener un máximo de 40 páginas (sin anexos) y debe tener la estructura siguiente:

1. Introducción incluyendo

    a. La descripción del objeto de la metaevaluación.

    b. La descripción de la metodología aplicada para la metaevaluación.

2. Metaevaluación de las evaluaciones,

    a. Un subcapítulo para cada evaluación evaluada incluyendo una corta descripción del programa evaluado (un párrafo), una corta descripción de la modalidad aplicada de implementación y apoyo a la evaluación[35] (máximo media página) y la calificación cuantitativa (con comentarios) de la evaluación y los respectivos términos de referencia según los criterios pre-establecidos. Cada subcapítulo debe cerrar con un resumen de las fortalezas y debilidades de la evaluación, con énfasis en la calidad general de la evaluación y el proceso evaluativo.

    b. Un capítulo de una apreciación general de las evaluaciones con una sinopsis de la calificación de las evaluaciones analizadas por la metaevaluación. El capítulo deberá incluir la descripción de patrones

[35] Bajo modalidad se entiende quién realizó la evaluación (personas contratadas, personas de las mismas instituciones, estudiantes etc.) y como fue apoyada la evaluación por FOCEVAL (aprender haciendo, padrinos = expertos que daban asesoría, etc.)

y tendencias generales, calificando la calidad general de las evaluaciones y los factores centrales que determinan la calidad de las evaluaciones y los procesos evaluativos.

3. Recomendaciones derivadas de la metaevaluación. Las evaluaciones deben referirse a la mejora de la calidad de las evaluaciones y de los términos de referencia, los procesos evaluativos y – si posible – la utilidad de las evaluaciones

4. Anexos

   - Términos de referencia de la metaevaluación
   - Una tabla que esquematiza la valoración cuantitativa de las evaluaciones y términos de referencia según cada uno de los criterios pre-establecidos
   - Documentos utilizados (informes de evaluaciones, términos de referencia y otros).

La metaevaluación se realizará en la oficina del contratista en base de los documentos enviados al contratista.

Un primer borrador del informe debe ser entregado hasta el día 30 de noviembre del 2016. El informe final debería incluir los comentarios realizados por el equipo técnico de FOCEVAL, y debe ser enviado hasta el día 19 de diciembre del 2016.

4. Responsabilidades

4a) FOCEVAL:

Como parte contratante, los responsables del proyecto FOCEVAL proporcionarán los documentos de las evaluaciones y darán seguimiento a la metaevaluación realizada por la persona experta.

4b) Persona experta:

Su responsabilidad principal es realizar una metaevaluación que servirá para valorar el indicador mencionado en párrafos anteriores, igualmente los resultados de la metaevaluación servirán como insumos para los talleres planeados y como base de las recomendaciones para futuros procesos de evaluación.

5. Perfil de la Persona Experta

A continuación se describen la cualificación y características que deberá reunir la persona experta contratada:

- Experiencia de al menos 5 años en el diseño, gestión, coordinación y ejecución de evaluaciones, preferentemente de programas o proyectos de ámbito nacional o regional.
- Experiencia de al menos 3 años en la formulación de términos de referencia.
- Tener conocimientos sobre la metaevaluación, idealmente tener experiencia en realizar metaevaluaciones en el ámbito nacional o regional.
- Tener conocimientos sobre los estándares internacionales de evaluación.
- Muy buen dominio del idioma español.

6. Especificación de insumos

La asistencia técnica será requerida en el periodo del 15 de setiembre del 2016 hasta el 19 de diciembre del 2016, para un total de hasta 14 días de trabajo. El borrador del informe debe ser presentado hasta el 30 de noviembre del 2016 y el informe final hasta el 19 de diciembre del 2016.

El número máximo de días de trabajo previsto para cada actividad es el siguiente:

Actividad	Días de trabajo hasta:
Metaevaluación	12
Adaptación del informe según la revisión por el equipo FOCEVAL	2
TOTAL días de trabajo hasta	**14**

7. Presentación de ofertas

Los/as candidatos/as deberán enviar sus ofertas por correo electrónico a Juan Sanz (juan.sanz@deval.org) y Ana Isabel Bruda (Isabel.bruda@deval.org), indicando en el Asunto "Experto en metaevaluación", antes del **10 de setiembre 2016 a las 7 p.m**., horario de Alemania.

Las ofertas deberán incluir

- El CV de la persona experta que demuestre que cumple con el perfil buscado.
- Una prueba escrita en español de máximo 1 página DIN A4 en la que se describa el uso de la metaevaluación, su relación con los estándares internacionales de evaluación y los retos en su realización.
- Una prueba escrita en español de máximo 1 página DIN A4 en la que se describa el rol de la triangulación para la calidad metodológica de evaluaciones.
- El informe final de una evaluación o metaevaluación coordinada/ejecutada por el oferente
- Precio del servicio de la asistencia técnica en euros.

Se valorará las ofertas según el perfil buscado (punto 5), la calidad de los documentos entregados y de las pruebas escritas. Se adjudicará el contrato teniendo en cuenta la correspondencia con el perfil buscado, la calidad de las pruebas escritas y la oferta económica.

5.2. Documentos utilizados

En este apartado se exponen los documentos utilizados como base para esta metaevaluación (informes de evaluaciones, términos de referencia y otros), que fueron recopilados y enviados por DEval.

	TÍTULO	TÍTULO CORTO	MARCO DE ELABORACIÓN	Documentos disponibles		
				Informe	TdR	Anexos
1.	Evaluación de Impacto del "Distrito de Riego Arenal-Tempisque", de Servicio Nacional de Áreas Subterráneas, Riego y Avenamiento (SENARA).	**DRAT**	FOCEVAL I	X		
2.	Evaluación del proyecto "Apoyo a la Gestión pedagógica a Centros educativos de calidad con orientación inclusiva", del Ministerio de Educación Pública (MEP).	**EDU-INCL**	FOCEVAL I	X		
3.	Evaluación del proyecto interinstitucional "Germinadora de Desarrollo Organizacional Empresarial, Asociativo y Comunitario".	**Germinadora**	FOCEVAL I	X		
4.	Evaluación del "Diseño Metodológico y Proceso de Elaboración del PND 2011 – 2014", de MIDEPLAN.	**PND**	FOCEVAL I	X		
5.	Evaluación del programa interinstitucional "Red Nacional de Cuido y Desarrollo Infantil (REDCUDI)".	**REDCUDI**	FOCEVAL II	X	X	X
6.	Evaluación del programa "Centros de Educación y Nutrición, y Centros Infantiles de Atención Integral (CEN-CINAI)", del Ministerio de Salud.	**CEN-CINAI**	FOCEVAL II	X	X	
7.	Evaluación del programa "Empléate", del Ministerio de Trabajo y Seguridad Social.	**Empléate**	FOCEVAL II	X		
8.	Evaluación del "Programa de Atención de la Oficina de Atención y Protección a la Víctima de Delito", del Poder Judicial de Costa Rica.	**OAPVD**	MEPPD	X		X
9.	Evaluación de efectos generados por los programas sociales que ejecuta el IMAS, 2009	**IMAS**	MEPPD	X	X	X
10.	Evaluación del programa de Atención Prenatal en el Primer Nivel de Atención del Ministerio de Salud	**Prenatal**	MinSalud	X		
11.	Evaluación participativa de los servicios de prevención y atención del cáncer en Valle la Estrella, provincia de Limón, (Costa Rica, 2017)	**PyA-Cáncer**	CCSS	X		X

5.3 Tabla sintética de valoración cuantitativa de informes de evaluación y TdR

CÓDIGO	CRITERIO / PREGUNTAS	SUBCRITERIO / DIMENSIONES / CATEGORIAS DE RESPUESTA	EVALUACIONES										
			DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA CÁNCER
			1	2	3	4	5	6	7	8	9	10	11
		PUNTUACION TOTAL ➔	86,8	69,1	78,0	73,0	67,5	72,1	71,7	71,6	78,0	58,1	73,1
1.	**Imparcialidad e Independencia**	**Calificación (texto)**	*10*	*7,5*	*10*	*10*	*7,5*	*7,5*	*7,5*	*7,5*	*10*	*5*	*5*
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	1	1	1	1	1	1	1	1	1	1
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	5	5	5	5	5	5	5	5	5	5	5
		Evaluación externa (100% externo a instituciones evaluadas)	10	10	10	10	10	10	10	10	10	10	10
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	1	1	1	1	1	1	1	1	1	1
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	5	5	5	5	5	5	5	5	5	5	5
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos)	10	10	10	10	10	10	10	10	10	10	10
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	*5,5*	*10*	*7,5*	*6,5*	*6*	*9*	*5,5*	*8,8*	*4,6*	*7,8*	*10*
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos? **Fases**: planificación e implementación.	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No hay participación	1	1	1	1	1	1	1	1	1	1	1
		Participan escasa o desigualmente en planificación o implementación	5	5	5	5	5	5	5	5	5	5	5
		Participan activamente en la planificación y en la implementación	10	10	10	10	10	10	10	10	10	10	10
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Las preguntas ni son claras ni adaptadas	1	1	1	1	1	1	1	1	1	1	1
		Las preguntas o son claras o adaptadas	5	5	5	5	5	5	5	5	5	5	5
		Las preguntas son claras y adaptadas	10	10	10	10	10	10	10	10	10	10	10
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	1	1	1	1	1	1	1	1	1	1
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	5	5	5	5	5	5	5	5	5	5
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	10	10	10	10	10	10	10	10	10	10
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No hubo un taller de inicio con stakeholders para facilitar su participación	1	1	1	1	1	1	1	1	1	1	1
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	5	5	5	5	5	5	5	5	5	5
		Hubo un taller de inicio con stakeholders	10	10	10	10	10	10	10	10	10	10	10
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No son accesibles para el público (ni Informe ni resumen)	1	1	1	1	1	1	1	1	1	1	1
		Son accesibles para el público (Resumen o extracto)	5	5	5	5	5	5	5	5	5	5	5
		Son accesibles para el público (Informe y resumen)	10	10	10	10	10	10	10	10	10	10	10
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*46,1*	*27,3*	*37,5*	*32,5*	*30,0*	*33,9*	*34,7*	*39,3*	*40,1*	*32,6*	*31,0*
3.1	**Metas, objetos y modelo de cambio**		*9,3*	*5,1*	*9,3*	*5,3*	*7,1*	*6,4*	*6,0*	*8,6*	*5,4*	*5,3*	*8,0*
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	1	1	1	1	1	1	1	1	1	1
		Se describe el objetivo de la evaluación de manera imprecisa	5	5	5	5	5	5	5	5	5	5	5
		Se describe de manera precisa el objetivo	10	10	10	10	10	10	10	10	10	10	10
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se describe el objeto de evaluación	1	1	1	1	1	1	1	1	1	1	1
		Se describe de manera imprecisa el objeto evaluado (>50% de items contemplados) o no contempla aspectos relevantes	5	5	5	5	5	5	5	5	5	5	5
		Se describe con precisión (al menos los items contemplados)	10	10	10	10	10	10	10	10	10	10	10
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se realizó un análisis de stakeholders	1	1	1	1	1	1	1	1	1	1	1
		Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.)	5	5	5	5	5	5	5	5	5	5	5
		Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	10	10	10	10	10	10	10	10	10	10	10

CÓDIGO	CRITERIO / PREGUNTAS	SUBCRITERIO / DIMENSIONES / CATEGORIAS DE RESPUESTA	EVALUACIONES										
			DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA CÁNCER
			1	2	3	4	5	6	7	8	9	10	11
		PUNTUACION TOTAL →	86,8	69,1	78,0	73,0	67,5	72,1	71,7	71,6	78,0	58,1	73,1
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se elaboró	1	1	1	1	1	1	1	1	1	1	1
		Se elaboró pero no se usa en la evaluación o no es concluyente	5	5	5	5	5	5	5	5	5	5	5
		Se elaboró, se usa en la evaluación y es concluyente	10	10	10	10	10	10	10	10	10	10	10
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se formularon	1	1	1	1	1	1	1	1	1	1	1
		Se formularon pero no son concluyentes	5	5	5	5	5	5	5	5	5	5	5
		Se formularon y son concluyentes	10	10	10	10	10	10	10	10	10	10	10
	¿Se diferencia entre los diferentes niveles de impacto?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se diferencia entre niveles de impacto	1	1	1	1	1	1	1	1	1	1	1
		Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)	5	5	5	5	5	5	5	5	5	5	5
		Se diferencia con claridad los diferentes niveles de impacto	10	10	10	10	10	10	10	10	10	10	10
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se describe el contexto de la evaluación	1	1	1	1	1	1	1	1	1	1	1
		Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención	5	5	5	5	5	5	5	5	5	5	5
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	10	10	10	10	10	10	10	10	10	10	10
3.2	**Diseño y métodos**	**Calificación (texto)**	*8,2*	*7,2*	*8,0*	*8,8*	*7,2*	*6,2*	*6,2*	*10,0*	*8,2*	*6,5*	*8,8*
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (*Los TdR no explicitan los intereses de la evaluación*)	0	0	0	0	0	0	0	0	0	0	0
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	1	1	1	1	1	1	1	1	1	1
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	5	5	5	5	5	5	5	5	5	5	5
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	10	10	10	10	10	10	10	10	10	10	10
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se explicitan ni justifican el diseño y los métodos aplicados	1	1	1	1	1	1	1	1	1	1	1
		Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados	5	5	5	5	5	5	5	5	5	5	5
		Se explicitan y justifican suficientemente el diseño y los métodos aplicados	10	10	10	10	10	10	10	10	10	10	10
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se consideró lo contrafactual	1	1	1	1	1	1	1	1	1	1	1
		Se consideró tangencialmente lo contrafactual o no se especifica suficientemente cómo se consideró	5	5	5	5	5	5	5	5	5	5	5
		Se consideró lo contrafactual y se especificó cómo	10	10	10	10	10	10	10	10	10	10	10
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Los métodos aplicados no cubren aspectos cuanti ni cuali	1	1	1	1	1	1	1	1	1	1	1
		Los métodos aplicados cubren aspectos cuanti o cuali	5	5	5	5	5	5	5	5	5	5	5
		Los métodos aplicados cubren aspectos cuanti y cuali	10	10	10	10	10	10	10	10	10	10	10
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Se aplicó sólo un método	1	1	1	1	1	1	1	1	1	1	1
		Se aplicaron dos métodos	5	5	5	5	5	5	5	5	5	5	5
		Se aplicaron una pluralidad de métodos (más de dos)	10	10	10	10	10	10	10	10	10	10	10
3.3	**Procedimiento de selección**	**Calificación (texto)**	*10,0*	*5,0*	*5,0*	*5,0*	*5,0*	*5,0*	*7,5*	*5,5*	*10,0*	*10,0*	
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se explicitan los procedimientos ni criterios de selección	1	1	1	1	1	1	1	1	1	1	1
		Se explicitan en algunos casos (técnicas)	5	5	5	5	5	5	5	5	5	5	5
		Se explicitan los procedimientos y criterios de selección	10	10	10	10	10	10	10	10	10	10	10
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? **Se entiende que aplica esta pregunta para las Encuestas**	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		La muestra no es adecuada	1	1	1	1	1	1	1	1	1	1	1
		La muestra no es adecuada o el error no es aceptable	5	5	5	5	5	5	5	5	5	5	5
		La muestra sí es adecuada y el error aceptable	10	10	10	10	10	10	10	10	10	10	10

CÓDIGO	CRITERIO / PREGUNTAS	SUBCRITERIO / DIMENSIONES / CATEGORIAS DE RESPUESTA	EVALUACIONES										
			DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA CÁNCER
			1	2	3	4	5	6	7	8	9	10	11
		PUNTUACION TOTAL ➔	86,8	69,1	78,0	73,0	67,5	72,1	71,7	71,6	78,0	58,1	73,1
3.4	**Análisis de datos**	**Calificación (texto)**	*10,0*	*3,0*	*6,3*	*5,5*	*5,3*	*6,3*	*7,5*	*7,5*	*6,5*	*5,5*	*4,0*
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se triangulan datos cuanti-cuali	1	1	1	1	1	1	1	1	1	1	1
		Se triangulan datos cuanti-cuali pero no de manera generalizada	5	5	5	5	5	5	5	5	5	5	5
		Se triangulan datos cuanti-cuali de manera generalizada	10	10	10	10	10	10	10	10	10	10	10
3.4.2	¿Se cuantifican la respuesta siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados…?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se cuantifican las respuestas cuando es posible	1	1	1	1	1	1	1	1	1	1	1
		Se cuantifican las respuestas a veces cuando es posible	5	5	5	5	5	5	5	5	5	5	5
		Se cuantifican las respuestas de manera generalizada cuando es posible	10	10	10	10	10	10	10	10	10	10	10
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se usaron datos del monitoreo	1	1	1	1	1	1	1	1	1	1	1
		Se usaron datos del monitoreo de manera insuficiente	5	5	5	5	5	5	5	5	5	5	5
		Se usaron datos del monitoreo de manera suficiente	10	10	10	10	10	10	10	10	10	10	10
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se recurrió a fuentes externas	1	1	1	1	1	1	1	1	1	1	1
		Se recurrió a fuentes externas de manera insuficiente	5	5	5	5	5	5	5	5	5	5	5
		Se recurrió a fuentes externas de manera suficiente	10	10	10	10	10	10	10	10	10	10	10
3.5	**Criterios de evaluación**	**Calificación (texto)**	*8,6*	*7,0*	*9,0*	*8,0*	*5,4*	*10,0*	*7,5*	*7,8*	*10,0*	*5,3*	*5,2*
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa												
a)	es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No lo expone	1	1	1	1	1	1	1	1	1	1	1
		Lo expone pero no de forma concluyente y comprensible	5	5	5	5	5	5	5	5	5	5	5
		Lo expone de forma concluyente y comprensible	10	10	10	10	10	10	10	10	10	10	10
b)	logró las metas definidas	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No lo expone	1	1	1	1	1	1	1	1	1	1	1
		Lo expone pero no de forma concluyente y comprensible	5	5	5	5	5	5	5	5	5	5	5
		Lo expone de forma concluyente y comprensible	10	10	10	10	10	10	10	10	10	10	10
c)	es eficiente	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No lo expone	1	1	1	1	1	1	1	1	1	1	1
		Lo expone pero no de forma concluyente y comprensible	5	5	5	5	5	5	5	5	5	5	5
		Lo expone de forma concluyente y comprensible	10	10	10	10	10	10	10	10	10	10	10
d)	causó impactos (outcome e impacto) deseados y no deseados.	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No lo expone	1	1	1	1	1	1	1	1	1	1	1
		Lo expone pero no de forma concluyente y comprensible	5	5	5	5	5	5	5	5	5	5	5
		Lo expone de forma concluyente y comprensible	10	10	10	10	10	10	10	10	10	10	10
e)	causó impactos duraderos	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No lo expone	1	1	1	1	1	1	1	1	1	1	1
		Lo expone pero no de forma concluyente y comprensible	5	5	5	5	5	5	5	5	5	5	5
		Lo expone de forma concluyente y comprensible	10	10	10	10	10	10	10	10	10	10	10
f)	cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No lo expone	1	1	1	1	1	1	1	1	1	1	1
		Lo expone pero no de forma concluyente y comprensible	5	5	5	5	5	5	5	5	5	5	5
		Lo expone de forma concluyente y comprensible	10	10	10	10	10	10	10	10	10	10	10
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	1	1	1	1	1	1	1	1	1	1
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	5	5	5	5	5	5	5	5	5	5	5
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	10	10	10	10	10	10	10	10	10	10

CÓDIGO	CRITERIO / PREGUNTAS	SUBCRITERIO / DIMENSIONES / CATEGORIAS DE RESPUESTA	DRAT	EDU INCL	Germinadora	PND	5 REDCUDI	CEN-CINAI	EMPLEATE	8 OAPVD	IMAS	PRENATAL	PyA CÁNCER
			1	2	3	4	5	6	7	8	9	10	11
4.	**Calidad del informe y utilidad**	**Calificación (texto)**	*18,8*	*18,8*	*16,5*	*18,8*	*18,8*	*17,8*	*17,5*	*11,8*	*17,8*	*7,5*	*18,8*
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	0	0	0	0	0	0	0	0	0	0
		No responde adecuadamente a ninguna pregunta	1	1	1	1	1	1	1	1	1	1	1
		No responde adecuadamente a alguna/s pregunta/s	5	5	5	5	5	5	5	5	5	5	5
		Responde adecuadamente a todas las preguntas	10	10	10	10	10	10	10	10	10	10	10
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se describen puntos fuertes y débiles	1	1	1	1	1	1	1	1	1	1	1
		Se describen principalmente puntos fuertes o débiles	5	5	5	5	5	5	5	5	5	5	5
		Se describen balanceadamente puntos fuertes y débiles	10	10	10	10	10	10	10	10	10	10	10
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	1	1	1	1	1	1	1	1	1	1
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	5	5	5	5	5	5	5	5	5	5
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	10	10	10	10	10	10	10	10	10	10	10
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se especifica	1	1	1	1	1	1	1	1	1	1	1
		Se especifica en ocasiones	5	5	5	5	5	5	5	5	5	5	5
		Se especifica en general	10	10	10	10	10	10	10	10	10	10	10
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Las conclusiones no están basadas en los resultados	1	1	1	1	1	1	1	1	1	1	1
		Alguna o algunas conclusiones no están basadas en resultados	5	5	5	5	5	5	5	5	5	5	5
		Todas las conclusiones están basadas en los resultados	10	10	10	10	10	10	10	10	10	10	10
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se llega ni de forma lógica ni comprensible	1	1	1	1	1	1	1	1	1	1	1
		Se llega de forma lógica o comprensible pero no las dos	5	5	5	5	5	5	5	5	5	5	5
		Se llega de forma lógica y comprensible	10	10	10	10	10	10	10	10	10	10	10
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No son relevantes ni específicas	1	1	1	1	1	1	1	1	1	1	1
		Son relevantes o específicas pero no las dos	5	5	5	5	5	5	5	5	5	5	5
		Son relevantes y específicas	10	10	10	10	10	10	10	10	10	10	10
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		El Informe no contiene un resumen	1	1	1	1	1	1	1	1	1	1	1
		El informe contiene un resumen estructurado	5	5	5	5	5	5	5	5	5	5	5
		El informe contiene un resumen estructurado, bien legible/comprensible	10	10	10	10	10	10	10	10	10	10	10
5.	**Exhaustividad y adecuación de la documentación**	**Calificación (texto)**	*6,5*	*5,5*	*6,5*	*5,3*	*5,3*	*4,0*	*6,5*	*4,3*	*5,5*	*5,3*	*8,3*
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		Los TdR no son accesibles en anexos	1	1	1	1	1	1	1	1	1	1	1
		Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	5	5	5	5	5	5	5	5	5	5
		Los TdR son accesibles en la lengua común de trabajo o en una traducción	10	10	10	10	10	10	10	10	10	10	10
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se incluyen en anexos	1	1	1	1	1	1	1	1	1	1	1
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	5	5	5	5	5	5	5	5	5	5	5
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	10	10	10	10	10	10	10	10	10	10
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		No se incluyen en anexos (guías y fuentes)	1	1	1	1	1	1	1	1	1	1	1
		Se incluyen algunos en anexos (guías o fuentes)	5	5	5	5	5	5	5	5	5	5	5
		Se incluyen en anexos (guías y fuentes)	10	10	10	10	10	10	10	10	10	10	10
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? Elementos que faciliten su lectura: gráficos, tablas y resúmenes de capítulos.	Ninguna / No aplica	0	0	0	0	0	0	0	0	0	0	0
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	1	1	1	1	1	1	1	1	1	1
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	5	5	5	5	5	5	5	5	5	5	5
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	10	10	10	10	10	10	10	10	10	10	10
		PUNTUACIÓN TOTAL	**86,8**	**69,1**	**78,0**	**73,0**	**67,5**	**72,1**	**71,7**	**71,6**	**78,0**	**58,1**	**73,1**

5.4 Ficha de registro de evaluaciones de la metaevaluación

4.4.1 Evaluación 01: DRAT

0.1	**Nombre del proyecto / programa evaluado**	**DISTRITO DE RIEGO ARENAL TEMPISQUE (DRAT) EN LA REGIÓN DE CHOROTEGA (1981-2011)**
0.2	**Institución que implementa (p.ej. Ministerio)**	SENARA Servicio Nacional de Aguas, Riego y Avenimiento (Páginas 17-18 y Página 22 nota al pie 3) http://www.fao.org/nr/water/aquastat/countries_regions/cri/indexesp.stm Junta Coordinadora del Distrito de Riego Arenal Tempisque (Páginas 17-18) INDER Instituto de Desarrollo Rural (antiguo IDA Instituto de Desarrollo Agrario) e ICE Instituto Costarricense de Electricidad (Página 50)
0.3	**Evaluadores**	Programa de Investigación en Desarrollo, Economía y Ambiente División de Investigación y Desarrollo, CATIE (Centro Agronómico Tropical de Investigación y Enseñanza). Coordinación de Juan Robalino. El equipo evaluador estuvo formado por Gloriana Lang, Chelsia Moraes, Catalina Sandoval, Luis Vargas y Laura Villalobos.
0.4	**Monto del proyecto/ programa**	Etapa I: US$ 20.6 millones. Etapa II: US$ 40.8 millones. Etapa III: US$ 3 millones. El monto total del proyecto asciende a US$ 64.4 millones.
0.5	**Monto para la evaluación**	No especificado en el Informe final de evaluación.

0.6 Descripción del proyecto/programa evaluado (un párrafo)
Se trata de un programa público tendente a promover el riego en las actividades de desarrollo agropecuario y rural del país. En concreto el proyecto del Distrito de Riego Arenal Tempisque (DRAT) fue implementado en 1978 con el objetivo de poner bajo desarrollo agrícola con riego tierras en la Región Chorotega y así fomentar el desarrollo agropecuario mediante la agricultura intensiva; buscando incrementar la producción de alimentos básicos. Además, se buscó mejorar la situación socioeconómica de la región por medio de la creación de nuevas fuentes de empleo y el aumento en el ingreso de los productores.
0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL (máximo media página)
El trabajo analizado consistió en una evaluación estratégica de Gobierno y fue realizada en el marco del Programa de Fortalecimiento de Capacidades en Evaluación (FOCEVAL) que se ejecuta en conjunto por parte de la Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ) y el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) por encargo del Ministerio Alemán de Cooperación Económica y Desarrollo (BMZ). La evaluación fue dirigida por un comité revisor conformado por miembros de GIZ Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (Jana Rauch), MIDEPLAN (Mario Gonzalez, Mainor Madrigal), SENARA (Juan Carlos Mora) y SEPSA (Ghiselle Rodríguez y Miriam Valverde).

0.8 Cualificación cuantitativa (con comentarios) según criterios			
86,8	**10**	***Imparcialidad e independencia:*** *[10 puntos]*	Se trata de una evaluación desarrollada por un equipo técnico externo a la organización que implementa el programa. Asimismo, las fuentes de información seleccionadas garantizaban la inclusión de pluralidad de perspectivas en el proceso evaluativo. Atendiendo a esto, se puede convenir que la evaluación garantizaba su imparcialidad e independencia.
	5,5	***Equidad, participación y transparencia:*** *[10 puntos]*	A nivel de documentación hay que indicar que si bien el Informe final de evaluación es accesible al público, no se ha facilitado al equipo que desarrolla esta meta evaluación los TdR. En relación a la participación, habría sido necesario implicar a otros grupos de interés en el proceso de planificación, y a los stackeholders de forma previa al lanzamiento de la evaluación.
	46,1	***Credibilidad y exactitud:*** *[50 puntos]*	El diseño metodológico de esta evaluación permite atribuirle un importante grado de exactitud y credibilidad. Tanto la planificación previa estableciendo una cadena de resultados, como la definición de las preguntas de evaluación y las hipótesis de procesos e impactos, permiten garantizar el desarrollo de un trabajo evaluativo de calidad. Tan sólo se aprecia como carencia el uso de un único método de análisis cualitativo (entrevistas en profundidad). Si bien se trata de una herramienta que aporta una información de calidad, habría resultado muy interesante complementarla con algún tipo de análisis donde se recabara la información con una perspectiva grupal (Grupos focales o estudios Delphi)

	18,8	***Calidad del informe y utilidad:*** *[20 puntos]*	El análisis de los resultados efectuado por el equipo evaluador, así como la redacción de conclusiones y recomendaciones hacen que el Informe final sea un documento de evaluación de bastante calidad. Al mismo tiempo, hay que indicar que tanto la redacción del informe como su estructura lo dotan de un elevado nivel de utilidad. La no disposición de los TdR es lo que impide alcanzar una nota de 10 en este apartado.
	6,5	***Exhaustividad y adecuación de la documentación:*** *[10 puntos]*	Si bien de forma global la documentación disponible es bastante exhausta y muy adecuada para los requerimientos que debe tener un informe final de evaluación, se echa en falta el documento de los TdR, así como información relativa al cronograma de trabajo y el presupuesto de la evaluación.

0.9 Resumen de fortalezas de la evaluación.

- ⇨ El equipo evaluador ha garantizado la independencia y pluralidad de perspectivas en el desarrollo del trabajo de evaluación.
- ⇨ Se ha efectuado un trabajo de triangulación excelente, obteniendo información relevante tanto del ámbito cualitativo como cuantitativo. Al mismo tiempo, la triangulación ha permitido cotejar información obtenida a partir de distintas fuentes, pero relativa a un mismo aspecto analizado.
- ⇨ El uso de técnicas de análisis cuantitativo que contemplan el diseño y aplicación de un contra-factual, supone contar con la posibilidad de poder inferir los resultados e impactos obtenidos a la existencia del programa de forma exclusiva.
- ⇨ El enfoque con el que se ha realizado el análisis de los resultados obtenidos durante el desarrollo del trabajo de campo no solo ha estado vinculado a obtener una valoración de los impactos que están derivándose del programa evaluado, sino también a la mejora en su implementación y a la mejora de futuras evaluaciones.
- ⇨ Tanto la estructura del informe como su redacción dotan al documento de un elevado nivel de calidad y utilidad.

0.10 Resumen de debilidades de la evaluación.

- ⇨ Cuando se evalúa un programa público con el alcance y la importancia que tiene el DRAT para un sector estratégico como es la agricultura, se hace necesario contar con una mayor diversidad de agentes clave en la planificación y diseño de la evaluación. De forma concreta la planificación debería incluir agentes involucrados en el diseño y gestión del programa, así como beneficiarios del mismo.
- ⇨ Aparece como carencia el no desarrollar algún tipo de actividad previa (jornada de sensibilización, presentación de la evaluación,…) a modo de instrumento que facilitara la participación de los stackeholders.
- ⇨ El uso de técnicas de análisis cuantitativo como el DID, el PSM y el control sintético, es complejo y no ha conseguido garantizar la existencia de un vínculo directo y exclusivo entre los resultados e impactos registrados y la existencia del programa. Quizás ante programas de la amplitud del DRAT se torna más interesante, a la vez que efectivo, desarrollar el estudio de los impactos a partir de modelos de evaluación que analizan la contribución del programa, frente a los de atribución utilizados en este trabajo.
- ⇨ El número de técnicas de recopilación de información cualitativa es limitado, en tanto que sólo se utiliza la entrevista en profundidad.
- ⇨ Existe una carencia relativa a información básica utilizada en el diseño metodológico. Si bien se hace referencia a un Estudio preliminar evaluativo (Imbach et al. 2013), se trata de una información de carácter relevante que no se incluye en el Informe final de evaluación.

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	Se trata de una evaluación externa en tanto que no se tiene constancia acerca de la participación de los miembros del equipo evaluador del Centro Agronómico Tropical de Investigación y Enseñanza (CATIE) en alguna de las instituciones evaluadas (SENARA).
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	5	
		Evaluación externa (100% externo a instituciones evaluadas)	**10**	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y análisis de datos y presentación de resultados	Ninguna / No aplica	0	Las fuentes de información elegidas por el equipo evaluador incluyeron a usuarios del servicio de riego (beneficiarios directos), a organizaciones que no son usuarias directas del servicio pero que desarrollan actividades dependientes de la presencia del riego (beneficiarios indirectos) y a instituciones públicas relacionadas o con inferencia en el Distrito de Riego Arenal Tempisque (DRAT). (Página 31)
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	5	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos).	**10**	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	***Calificación (numérico)***	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos? **Fases**: planificación e implementación.	Ninguna / No aplica / No hay información en los documentos consultados	0	En base a la información disponible en el informe no se involucró a grupos interesados en la planificación e implementación del proceso de evaluación.
		No hay participación	**1**	
		Participan escasa o desigualmente en planificación o implementación	5	
		Participan activamente en la planificación y en la implementación	10	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	Dentro del punto 4.2 Interrogantes evaluativas y sus alcances, se explicitan las diferentes preguntas de evaluación presentando una interrogante principal y un conjunto de interrogantes secundarias a través de las que dar respuesta a los objetivos de la evaluación. (Páginas 22-26)
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	5	
		Las preguntas son claras y adaptadas	**10**	
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica**	**0**	No se dispone de los TdR.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	Si bien no se realizó un taller de inicio, se utilizaron dos formatos orientadores para la realización de las entrevistas. (Página 31)
		No hubo un taller de inicio con stakeholders para facilitar su participación	**1**	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	
		Hubo un taller de inicio con stakeholders	10	
2.5		Ninguna / No aplica	0	El informe final de evaluación se encuentra disponible a través del siguiente link:

	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	No son accesibles para el público (ni Informe ni resumen)	1	https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/be0001a6-fed8-49a2-9ad4-6efe2fa6f283/DRAT%20Informe%20de%20evaluacion.Pdf?guest=true
		Son accesibles para el público (Resumen o extracto)	5	
		Son accesibles para el público (Informe y resumen)	**10**	
3.	**Credibilidad y exactitud**	**Calificación (texto)**	**Calificación (numérico)**	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**			
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	En el punto 4 Proceso de la evaluación y diseño, se describe explícitamente tanto el objetivo general como los objetivos específicos de la evaluación. (Página 16)
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	
		Se describe el objetivo de la evaluación de manera imprecisa	5	
		Se describe de manera precisa el objetivo	**10**	
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica	0	Dentro del punto 3 Descripción de la intervención, se efectúa una descripción precisa del objeto de evaluación. Se detalla información relativa a: (Páginas 15-21) • Zona de influencia, etapas de la intervención y volumen de la inversión en cada etapa. • Objetivos del programa (iniciales y reformulados). • Insumos, actividades y productos desglosados por etapas. • Efectos producidos, efectos esperados, y supuestos presentados por etapas.
		No se describe el objeto de evaluación	1	
		Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes	5	
		Se describe con precisión (al menos los ítems contemplados)	**10**	
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica	0	En el epígrafe 4.4.1.4 Fuentes de información, se presenta la información relativa a los grupos de interés que participaron en el trabajo de campo. Se indica que la participación de los actores informantes clave se determinó en base al mapeo de actores presentado en el Estudio preliminar evaluativo (Imbach et al. 2013). Página (31)
		No se realizó un análisis de stakeholders	1	
		Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.)	5	
		Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	**10**	
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica	0	En el informe de final de evaluación se indica que en el Estudio preliminar evaluativo (Imbach et al. 2013) se desarrolló un mapa que representa la teoría de cambio. Este mapa se incluye en Apéndice 1 del informe. A través del desarrollo de una cadena de resultados se definió la teoría de cambio del proyecto, siendo el primer objetivo del trabajo de campo validar el mapa causal definido. (Página 29)
		No se elaboró	1	
		Se elaboró pero no se usa en la evaluación o no es concluyente	5	
		Se elaboró, se usa en la evaluación y es concluyente	**10**	
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica	0	El trabajo de evaluación incluyó tanto la formulación de Hipótesis de supuestos como de Hipótesis de Impacto. A través del trabajo de campo cualitativo y de su cotejo con los resultados obtenidos en el análisis cuantitativo, se validaros los resultados de las Hipótesis de impacto planteadas. (Páginas 29-31)
		No se formularon	1	
		Se formularon pero no son concluyentes	5	
		Se formularon y son concluyentes	**10**	
	¿Se diferencia entre los diferentes niveles de impacto? Niveles de impacto:	Ninguna / No aplica	0	Desde su inicio el informe final de evaluación presenta los diferentes niveles (variables agrícolas, socioeconómicas y ambientales seleccionadas) así como factores (impactos ambientales, formación de organizaciones e integración de servicios públicos para la producción y comercialización), en los que se va a medir el impacto. Asimismo, las Hipótesis de Impacto establecidas diferencian entre distintos niveles de impacto. (Páginas 8 y 30).
		No se diferencia entre niveles de impacto	1	
		Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)	5	
		Se diferencia con claridad los diferentes niveles de impacto	**10**	
3.1.5		Ninguna / No aplica	0	Se describe el contexto de la evaluación y se relaciona con la teoría, aunque no de forma exhaustiva ni es un elemento central del informe.
		No se describe el contexto de la evaluación	1	

	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	**Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención**	**5**	
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	10	
3.2	**Diseño y métodos**	**Calificación (texto)**	**Calificación (numérico)**	**Observaciones cualitativas**
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (Los TdR no explicitan los intereses de la evaluación)	0	No se dispone de los TdR. Sin embargo, hay que indicar que el diseño metodológico de esta evaluación se corresponde de forma adecuada con los objetivos establecidos para la misma que se presentan en el informe final de evaluación.
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	5	
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	**10**	
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica	0	Dentro del punto 4 Proceso de la evaluación y diseño, se explicita y justifica con un nivel de detalle más que aceptable los métodos de evaluación aplicados (cualitativos y cuantitativos), las fuentes de información utilizadas y los instrumentos de análisis utilizados para procesar la información obtenida. (Páginas 22-43)
		No se explicitan ni justifican el diseño y los métodos aplicados	1	
		Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados	5	
		Se explicitan y justifican suficientemente el diseño y los métodos aplicados	**10**	
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	Ninguna / No aplica	0	Para la estimación del contra-factual, se utilizaron métodos cuantitativos tales como regresión, diferencias en diferencias y "matching". Asimismo, con el objetivo de explicar los posibles resultados, se complementó el análisis con el uso de métodos cualitativos. De forma concreta, el epígrafe 4.4.2.2 Métodos de análisis de datos, detalla el uso de las siguientes metodologías: Diferencias en Diferencias, Propensity Score Matching y Control sintético. (Páginas 8 y 35-38)
		No se consideró lo contrafactual	1	
		Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró	5	
		Se consideró lo contrafactual y se especificó cómo	**10**	
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica	0	Dentro del punto 4 Proceso de la evaluación y diseño, se presentan los métodos de análisis cuasi-experimentales que se complementaron con entrevistas cualitativas. (Páginas 22-43)
		Los métodos aplicados no cubren aspectos cuanti ni cuali	1	
		Los métodos aplicados cubren aspectos cuanti o cuali	5	
		Los métodos aplicados cubren aspectos cuanti y cuali	**10**	
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? A) entrevistas guiadas (número?); b) grupos focales (número?);c) encuestas (semi-) estructuradas (número?)	Ninguna / No aplica	0	En total se realizaron 28 entrevistas, siendo 14 a instituciones, 12 a beneficiarios directos y 2 a beneficiarios indirectos (Página 31). Además, una de las fuentes para el análisis cuantitativo es la información de las Encuestas de Hogares de Propósitos Múltiples (EHPM) realizadas cada año durante el período 2000-2009. (Página 33)
		Se aplicó sólo un método	**1**	
		Se aplicaron dos métodos	5	
		Se aplicaron una pluralidad de métodos (más de dos)	10	

	¿de qué forma (face-to-face, online, teléfono etc.);d) otros?..			
3.3	**Procedimiento de selección**	**Calificación (texto)**	**Calificación (numérico)**	**Observaciones cualitativas**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección Se explicitan en algunos casos (técnicas) **Se explicitan los procedimientos y criterios de selección**	0 1 5 **10**	Dentro del epígrafe 4.4.1.4 Fuentes de información, se explicitan los diferentes grupos participantes en las entrevistas. En este sentido, se diferencia entre actores beneficiarios (directos e indirectos) y actores institucionales. Asimismo se caracteriza a los beneficiarios directos según tipo de cultivo y extensión de las fincas, y se representan a los beneficiarios indirectos a través de una cooperativa y una empresa productora. Sin embargo, no se define el procedimiento para seleccionar a los 28 entrevistados de entre el total de potenciales participantes. Se vuelve a hacer referencia al Estudio preliminar evaluativo (Imbach et al. 2013) "*…se definió un grupo de actores de mayor relevancia o con mayor grado de relación con la intervención con base en el mapeo de actores presentado en el Estudio Preliminar…*" (Página 31).
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? **Se entiende que aplica esta pregunta para las Encuestas**	**Ninguna / No aplica** La muestra no es adecuada La muestra no es adecuada o el error no es aceptable La muestra sí es adecuada y el error aceptable	**0** 1 5 10	La evaluación analizada no contempló la realización de procesos de encuestación. Por lo tanto, no cabe hablar de ningún tipo de muestra seleccionada.
3.4	**Análisis de datos**	**Calificación (texto)**	**Calificación (numérico)**	**Observaciones cualitativas**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali Se triangulan datos cuanti-cuali pero no de manera generalizada **Se triangulan datos cuanti-cuali de manera generalizada**	0 1 5 **10**	4.4.1.3 Validación de los resultados Finalmente, se hizo una comparación entre la percepción de los efectos por parte de los actores y los resultados generados cuantitativamente. Esto permitió validar los análisis cuantitativos e identificar posibles explicaciones. (Página 31)
3.4.2	¿Se cuantifican la respuesta siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados…?	Ninguna / No aplica No se cuantifican las respuestas cuando es posible Se cuantifican las respuestas a veces cuando es posible **Se cuantifican las respuestas de manera generalizada cuando es posible**	0 1 5 **10**	A lo largo del punto 5 Resultados, se detallan y ponderan las respuestas de los stakeholders a las cuestiones de carácter cualitativo analizadas en las entrevistas, relativas a la verificación de las Hipótesis de procesos e Hipótesis de impactos. Aunque dentro de este punto del informe no existe una cuantificación o ponderación numérica de las respuestas, en el Apéndice 3 se presenta el porcentaje de respuesta según supuesto o interrogante, para ambos tipos de hipótesis. (Página 84)
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica No se usaron datos del monitoreo Se usaron datos del monitoreo de manera insuficiente **Se usaron datos del monitoreo de manera suficiente**	0 1 5 **10**	Se usaron series históricas de datos de monitoreo del Servicio Nacional de Aguas, Riego y Avenimiento SENARA. (Páginas 2 y 26)
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica No se recurrió a fuentes externas Se recurrió a fuentes externas de manera insuficiente **Se recurrió a fuentes externas de manera suficiente**	0 1 5 **10**	Por un lado, para el análisis de calidad del agua se usaron datos del Servicio Nacional de Aguas, Riego y Avenimiento SENARA. Por otra parte, para obtener determinados indicadores cuantitativos se recogieron datos de los censos nacionales de población y vivienda, y de la Encuesta de Hogares de Propósitos Múltiples que levanta el Instituto Nacional de Estadística y Censo. Asimismo, para obtener indicadores de producción agrícola se recurrió a información del Consejo Nacional de Producción y de la Liga Agrícola Industrial de la Caña de Azúcar.(Páginas 2, 31, 33 y 34)
3.5	**Criterios de evaluación**	**Calificación (texto)**	**Calificación (numérico)**	**Observaciones cualitativas**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			

	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	Tanto en el punto 2 Contexto y origen de la intervención evaluada, como en el 3 Descripción de la intervención, el informe deja clara la importancia del proyecto para el desarrollo económico de la región y del país. Se explicita de forma clara la relevancia del regadío en la producción agrícola y la pertinencia de implementar una política pública de este tipo para mejorar la calidad de vida de la población beneficiaria.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	5	
		Lo expone de forma concluyente y comprensible	**10**	
	b) logró las metas definidas	Ninguna / No aplica	0	Tanto desde un punto de vista cualitativo como cuantitativo, dentro del punto 5 relativo a los Resultados el Informe final expone de forma detallada los resultados e impactos del programa, dando respuesta a las preguntas de evaluación e hipótesis formuladas (de procesos y de impactos) Esto supone que el análisis relativo al logro de metas definidas es claro y concluyente.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	c) es eficiente	Ninguna / No aplica	0	Dentro del documento final no se encuentra información concluyente relativa a un análisis de eficiencia del programa. En este sentido, tan sólo en el análisis cualitativo se plantea por parte de algunos stakeholders la necesidad de implementar medidas para buscar una mayor eficiencia de la gestión del agua. (Páginas 62 y 70)
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	
		Lo expone de forma concluyente y comprensible	10	
	d) causó impactos (outcome e impacto) deseados y no deseados.	Ninguna / No aplica	0	El Informe final expone de forma concluyente los resultados del programa: En este sentido hay que indicar que, si bien desde el análisis cualitativo se indica que en muchos aspectos el programa tuvo impactos positivos en la población beneficiaria, los resultados obtenidos a partir del análisis cuantitativo no reflejan la misma situación. Así, desde un punto de vista cuantitativo, si bien para algunos ítems se obtienen impactos positivos del programa, también se indica para la totalidad de variables analizadas que los resultados presentan un margen de error que incluye la posibilidad de que pudo no haber impacto.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	e) causó impactos duraderos	Ninguna / No aplica	0	Tal y como se observa en el apartado anterior, no se puede hablar de la existencia de impactos duraderos en tanto que estadísticamente, no se puede eliminar la posibilidad de ausencia de impacto debido a que el cero está también dentro del margen de error.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	El proyecto del Distrito de Riego Arenal Tempisque cuenta con un sistema de monitoreo gestionado por el Servicio Nacional de Aguas, Riego y Avenimiento (SENARA). Durante el trabajo de evaluación se usaron datos concretos y series históricas de datos de monitoreo, para obtener determinados indicadores de carácter cualitativo como el de calidad del agua. (Páginas 2, 26 y 31)
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	Si bien a partir de la información cualitativa proporcionada por los actores participantes en la evaluación se podría convenir que los cambios observados se derivan del proyecto del Distrito de Riego Arenal Tempisque, el análisis cuantitativo desarrollado (Diferencias en Diferencias, Propensity Score Matching y Control sintético) no permiten afirmar que esto sea cierto de manera concluyente.
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	**5**	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas

4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	No se dispone de los TdR. En el informe se indica que las preguntas se obtienen del estudio pre-evaluativo (Imbach et al., 2013). Hecha esta salvedad, el informe da respuesta a las preguntas planteadas
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	El análisis de los resultados presentado en el punto 6 relativo a Conclusiones detalla tanto aspectos fuertes como débiles del programa, basados en la evidencia de los análisis cuantitativos y cualitativos realizados.
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	La información se presenta de forma que resulta sencillo distinguir entre reflexiones descriptivas y valorativas. En este sentido se pone de manifiesto claramente el trabajo de triangulación realizado por el equipo evaluador que suponía, en parte, cotejar la información cualitativa obtenida con los resultados del análisis cuantitativo.
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	**10**	
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	La fuente de información cualitativa (stakeholders) es de carácter primario y atendiendo a la caracterización de los participantes es confiable. Las fuentes de información cuantitativa son registros públicos de carácter primario y confiable: • Instituto Nacional de Estadística y Censos INEC (Censos Nacionales de Población y Vivienda y Encuestas de Hogares de Propósitos Múltiples) • Consejo Nacional de Producción (CNP) y Corporación Arrocera Nacional.
		No se especifica	1	
		Se especifica en ocasiones	5	
		Se especifica en general	**10**	
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica	0	Tanto los resultados obtenidos en la evaluación, como las conclusiones redactadas por el equipo evaluador, están estructurados para dar respuesta a la interrogante evaluativa principal y a las cinco interrogantes evaluativas secundarias que la articulan. Se sigue por tanto un proceso lógico para determinar conclusiones, vinculado a los resultados obtenidos.
		Las conclusiones no están basadas en los resultados	1	
		Alguna o algunas conclusiones no están basadas en resultados	5	
		Todas las conclusiones están basadas en los resultados	**10**	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	Tanto el Informe final como los resultados, conclusiones y recomendaciones se presentan de forma estructurada, lógica y comprensible. Sin embargo la estructura sigue la siguiente secuencia: Resultados > Conclusiones > Recomendaciones
		No se llega ni de forma lógica ni comprensible	1	
		Se llega de forma lógica o comprensible pero no las dos	5	
		Se llega de forma lógica y comprensible	**10**	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	El listado de recomendaciones propuesto es relevante y específico, tanto para la mejora del programa (incluyendo a la población beneficiaria) como para la mejora de futuras evaluaciones del mismo, o de programas similares.
		No son relevantes ni específicas	1	
		Son relevantes o específicas pero no las dos	5	
		Son relevantes y específicas	**10**	
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	A modo de preámbulo, el informe incluye un resumen ejecutivo que algo más de una página presenta las ideas clave de la evaluación y de sus resultados. Este resumen si bien es legible y comprensible, y presenta de forma sintética los aspectos claves de la evaluación.
		El Informe no contiene un resumen	1	
		El informe contiene un resumen estructurado	**5**	

		El informe contiene un resumen estructurado, bien legible/comprensible	10	No obstante debiera aportar más información de forma que el lector pueda conocer más detalle sobre los resultados y los límites de su validez, exponiendo una síntesis de las principales conclusiones y recomendaciones.
5.	**Exhaustividad y adecuación de la documentación**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica	0	No se dispone de los TdR.
		Los TdR no son accesibles en anexos	**1**	
		Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	
		Los TdR son accesibles en la lengua común de trabajo o en una traducción	10	
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	Se incluyen lugares, organizaciones visitadas y personas entrevistadas, respetando siempre la confidencialidad. No se incluye cronograma de trabajo de campo. (Apéndice 2. Listado de Entrevistados en Páginas 82-83)
		No se incluyen en anexos	1	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	**5**	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	En esta evaluación no se realizaron grupos focales ni encuestas. En relación a las entrevistas en profundidad, se incluyen tanto el listado de participantes como los guiones de entrevistas utilizados. (Apéndice 2. Listado de Entrevistados en Páginas 82-83) (Anexo 1. Guía de Preguntas a usuarios en Páginas 87-94) (Anexo 2. Guía de preguntas a actores institucionales en Páginas 95-100)
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	**10**	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	Ninguna / No aplica	0	A excepción de los resúmenes de los capítulos, el informe final de evaluación presenta una estructura más que adecuada para facilitar su lectura. Tanto la forma en la que se estructura el informe a nivel de capítulos, como la distribución de los epígrafes dentro de los mismos, facilitan su lectura y entendimiento. Asimismo, se presentan gráficos y tablas en un número suficiente. Incluyendo siempre una información clara, concisa y detallada.
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	5	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	**10**	

4.4.2 Evaluación 02: EDU-INCLUSIVA

0.1	**Nombre del proyecto / programa evaluado**	***PROYECTO: APOYO A LA GESTIÓN PEDAGÓGICA A CENTROS EDUCATIVOS DE CALIDAD CON ORIENTACIÓN INCLUSIVA***
0.2	**Institución que implementa**	*Dirección de Desarrollo Curricular del Ministerio de Educación Pública (MEP) de Costa Rica, desde la acción de la Comisión de Apoyo a la Educación Inclusiva (CAEI), integrada por miembros de todos los departamentos de dicha Dirección y del Centro Nacional de Recursos para la Educación Inclusiva (CENAREC)*
0.3	**Evaluadores**	*MIDEPLAN (Ministerio de Planificación Nacional y Política Económica): Marjorie Valerín, María José Vega* *MEP (Ministerio de Educación Pública): Mario Segura, Rafael Martínez* *GIZ (Agencia Alemana de Cooperación Internacional): Mayela Zúñiga* *Consultora: Susan Francis*
0.4	**Monto del proyecto/ programa**	*No especificado en Informe final de evaluación (Anexo 9 únicos datos financieros).* *En la página 37 se indica que según POA de la Dirección de Desarrollo Curricular para el 2011, se contó con un presupuesto de ¢38.159.200 millones para diversas actividades propias del proyecto. Y, en el propio Anexo 9 se detallan las "Actividades del Plan Operativo Anual vinculado al proyecto Dirección de Desarrollo Curricular 2011 " y aquellas que poseen presupuesto vinculado* ***Actividad*** — ***Presupuesto*** *Capacitación a funcionarios de los departamentos de Asesoría Pedagógica y Supervisores mediante el curso "Estrategias para avanzar hacia la Educación Inclusiva". — ¢22.659.200,00* *Capacitación a equipos de trabajo de los centros educativos del Proyecto "Apoyo a la gestión pedagógica de los centros educativos de calidad con orientación inclusiva" — ¢13.900.000,00* *Elaboración de un desplegable de apoyo a la gestión curricular de los centros educativos en materia de Educación Inclusiva. — ¢1.600.000,00*
0.5	**Monto para la evaluación**	*No especificado en Informe final de evaluación*
0.6	**Descripción del proyecto/programa evaluado**	*Este proyecto aborda la problemática de la exclusión del sistema educativo basándose en que se evidencia en un contexto caracterizado por la desarticulación entre el centro educativo y las necesidades del estudiantado, de las familias y de la comunidad, que genera educación desigual y excluyente para el estudiantado y lleva a peores registros de reprobación y a la deserción escolar, a tener menos oportunidades y a una menor calidad de vida; situaciones que no se conciben desde la esfera individual del alumno sino que operan desde el sistema educativo en su conjunto. Para combatir esta problemática propone una experiencia piloto en una muestra de centros educativos entre junio de 2009 y diciembre de 2012 (6 inician en 2009 y 4 se incorporan en 2011). Se trata de generar una transformación en estos centros que produzcan prácticas y políticas institucionales para impulsar una cultura institucional inclusiva, mediante la integración de la comunidad educativa (familias, estudiantado, personal docente, técnico y administrativo y otros) y a través de la aplicación de la promoción de los fundamentos de la Educación inclusiva mediante la aplicación del denominado Índice de Inclusión. Este índice, constituye una estrategia metodológica conformada por un conjunto de materiales diseñados para apoyar a las escuelas en el proceso de cambio hacia comunidades escolares colaborativas que promuevan en todo el alumnado altos niveles de logro; y se desarrolla en una serie de etapas. Se establecen siete etapas: (1) Construcción y aprobación del proyecto; (2) Concienciación; (3) Capacitación a equipos de trabajo; (4) Estudio del contexto escolar; (5) Seguimiento a la propuesta; (6) Acompañamiento pedagógico a las instituciones educativas; y (7) evaluación y la sistematización de la experiencia. Es la aplicación secuencial de estas etapas la que determina la aplicación del Índice de inclusión en los centros educativos (objetivo general) y de la que se espera que se deriven prácticas y políticas inclusivas, generando una cultura inclusiva (objetivos específicos). En definitiva, este proyecto se propone bajo la premisa que la educación inclusiva es un movimiento que se opone a cualquier forma de discriminación en el ámbito educativo, es una oportunidad y herramienta para dar respuesta a la diversidad de la población y que su promoción es una aportación a la educación pública de calidad. (Aguilar, 2013; citado en Pág.21).*
0.7	**Descripción de la modalidad de implementación/apoyo de FOCEVAL**	***Evaluación desarrollada con el apoyo de FOCEVAL bajo la modalidad de Aprender-haciendo, mediante una consultora nacional y la conformación de un grupo de trabajo MIDEPLAN-GIZ-MEP.*** *Este proyecto fue presentado por la CAEI como cierre de la intervención al concurso de Evaluaciones piloto que en noviembre – diciembre 2012 convocaron MIDEPLAN junto con la GIZ, en el marco del programa de FOCEVAL. Esto se plasmó en la conformación para tales efectos de un grupo de trabajo con la participación de representantes de MIDEPLAN quienes fungen como funcionarios de la Unidad de Acciones Estratégicas (UAE) del Área Evaluación y Seguimiento (AES), así como representantes de GIZ y del Ministerio de Educación Pública (MEP). Para su realización se contrató una consultora nacional experta en evaluación y educación.* *En la lectura del Informe se aprecian referencias explícitas a esta modalidad de apoyo, que también se referencian en anexos (Anexo 2) donde se explicitan las expectativas de participación en la evaluación de MIDEPLAN, FOCEVAL y GIZ vinculadas al desarrollo de la capacidad de evaluación.* *Entre los aspectos positivos que pueden destacarse de esta modalidad de apoyo en esta evaluación cabe señalar que el propio equipo evaluador destaca en su informe que el proceso de aprender-haciendo ha generado importantes y valiosos aprendizajes, entre los participantes, y que facilita "reconstruir conocimientos en prácticas reales".* *"… se considera valioso todo el aprendizaje logrado por el proceso evaluativo para todas las personas que participaron en él, la consideración del proceso de **aprender haciendo** establece oportunidades de reconstrucción de conocimientos en prácticas reales".* *Entre los aspectos negativos, condicionantes o limitaciones que ha supuesto su aplicación, en el informe se señala al respecto del proceso de aprender-haciendo que:* *"No obstante, se requieren bases previas para asegurar que no se convierta en un proceso extenso en el tiempo y, por tanto, de alta inversión"... " (Pág. 35)* *Se señala "la estrategia de articulación institucional para desarrollar la evaluación", como una de las dos principales limitaciones que tuvo este proceso evaluativo (junto a la identificación de la teoría de la intervención). A este respecto se apunta que la modalidad de aprender-haciendo: "requiere la previsión de las capacidades de aprendizaje y trabajo en equipo interinstitucional". Concretamente se apunta que:*

		"La modalidad de «Aprender haciendo» es un ejercicio muy complejo y costoso, lo cual implica altos niveles de compromiso, clarificación en los papeles y esfuerzos de coordinación importantes entre las instituciones participantes, para el desarrollo efectivo de evaluaciones sobre la gestión de las instituciones públicas" (Pág. 17 y 68). *Basta hacer una lectura de las lecciones aprendidas para comprender las tensiones que ha supuesto su aplicación en este proceso evaluativo. El hecho de que no basta con un perfil experto en evaluación (se necesita también amplio conocimiento del funcionamiento del sector público y de la política evaluada –Pág.17-) o de que la participación de representantes de la institución responsable del proyecto evaluado en el equipo evaluador genera tensiones que deben quedar resueltas.* *"En la planificación, la organización y la ejecución de procesos de evaluaciones públicas, la posibilidad de participación de miembros de la unidad ejecutora de la intervención que genera una serie de condiciones que deberán ser analizadas a priori: función que cumplen en el equipo evaluador, actividades que técnicamente están habilitadas para realizar, compromisos y productos de su participación en la evaluación. Además, es fundamental establecer en los procesos de negociación interinstitucional, los elementos de carácter ético y conceptual que implica integrar personal que participó de la ejecución de la intervención". (Pág. 16).*

0.8	**Cualificación cuantitativa (con comentarios) según criterios** *[Sobre 100]*	**69,1**	**7,5**	*Imparcialidad e independencia: [Sobre 10]*	En el Informe se presenta el hecho de que la propia institución evaluada participara en la realización de la evaluación pudo ser un factor limitante, sin embargo esto se contrarresta con la utilización de una pluralidad de técnicas y fuentes, que ayuda a tener una visión global y completa.
			10	*Equidad, participación y transparencia: [Sobre 10]*	Esta evaluación destaca especialmente por la implicación y participación de los diversos actores implicados, que son involucrados no sólo desde el rol convencional de suministradores de información (informantes clave), destacando en este sentido la participación activa de la CAEI en la fase de diseño e implementación. También en transparencia los resultados son satisfactorios, ya que los productos de la evaluación pueden ser consultados por cualquier persona a través de Internet, y las preguntas de evaluación son claras y permiten hacerse una idea adecuada de los focos donde se centra esta evaluación, aunque éstas hayan requerido un procesamiento respecto a las inicialmente formuladas según se indica en el informe. El único ítem que no aplicó fue la disposición de los TdR, ya que no estuvieron disponibles en esta metaevaluación.
			27,3	*Credibilidad y exactitud: [Sobre 10]*	Las puntuaciones en las que se desglosa este criterio resultan desiguales en esta evaluación. Hay aspectos que están consolidados, como el análisis de los stakeholders y su uso en la evaluación, la correspondencia del diseño mediante estudios de caso a los intereses de la evaluación, o la pluralidad de métodos empleados que cubren tanto lo cuantitativo como lo cualitativo, así como el tratamiento aceptable y concluyente que reciben los principales criterios de evaluación que le son de aplicación. Así, sus mejores registros en términos de credibilidad y exactitud se centran en el diseño y los métodos aplicados, así como en su tratamiento y respuesta a los criterios de evaluación contemplados. Por el contrario, las dimensiones en las que recibe una valoración intermedia están asociadas con la definición de metas, objetos y modelo de cambio, la especificación de los procedimientos de selección -a los que presta escasa atención-, y especialmente en la relativa al análisis de datos, en la que le penaliza especialmente el que no contase con datos del monitoreo del programa -no disponibles según se indica- ni recurriese a fuentes de datos externas para facilitar la contextualización y comparación.
			18,8	*Calidad del informe y utilidad: [Sobre 20]*	En este criterio la evaluación obtiene sus mejores registros, pues presenta la más alta valoración en la mayoría de ítems contemplados, siendo su principal debilidad la escasa atención que presta a las cuestiones de fiabilidad en los hallazgos o conclusiones presentados.
			10	*Exhaustividad y adecuación de la documentación: [Sobre 10]*	La exhaustividad y adecuación de la documentación destaca en cuanto que presenta una estructura adecuada y lectura apoyada en tablas y gráficos así como en la inclusión en anexos de los instrumentos aplicados para recabar información. No obstante, sus limitaciones se centran en no presentar anexos los TdR y en la mejorable especificación en anexos del trabajo de campo realizado.
0.9	**Resumen de fortalezas de la evaluación.**	• *Realiza un amplio trabajo de campo cualitativo.* • *Alta variedad de stakeholders contemplados y participantes como informantes clave.* • *La participación como un eje relevante del proceso, más allá de la mera consulta de agentes informantes clave, especialmente desarrollada con CAEI.* ➘ *Destaca la construcción participativa de preguntas y delimitación del encargo evaluativo con la CAEI.* • *El establecimiento de recomendaciones por actor supone una importante orientación práctica que debe facilitar su aplicación así como el establecimiento de un plan de acción para el seguimiento de la implementación de estas recomendaciones.* • *La operacionalización de los criterios evaluativos, el establecimiento de categorías de análisis y parámetros para su valoración, definiendo elementos, patrones y estructuras para cada una de ellas, así como su operacionalización en indicadores (Anexo 5) refuerza el diseño de la evaluación, y el tratamiento sistemático de la abundante información cualitativa recabada.* • *El esfuerzo por introducir verbatim entre los hallazgos con frases literales de las personas entrevistas, que contextualizan acertadamente aquello que se expone, especialmente relativos a la entrevista con la CAEI, pero también de entrevistas grupales y algún estudio de caso.*				
0.10	**Resumen de debilidades de la evaluación.**	• *En el diseño de la evaluación se apuntan enfoques o modelos e indicadores para los que se aprecia escaso o limitado tratamiento en el desarrollo de la evaluación.*			

		➘ *El modelo CIPP que anuncia como anclaje (contexto, insumos, procesos, productos), considerándose acertado, no se aprecia con claridad su uso más allá de la exposición de una "cadena de resultados" basada en sus componentes.* ➘ *Los indicadores del Índice de inclusión que se indican y se presentan en anexo no se utilizan para estudiar y comparar los casos (Centros educativos)* ➘ *Centra el uso de métodos y técnicas en los estudios de caso (que puede ser acertado teniendo en cuenta el objetivo explícito de la sistematización de la experiencia), si bien, no se incorporan ni anexan los estudios de caso de los que se nutre y las referencias a sus resultados son difusas o dificultan la trazabilidad de los hallazgos, conclusiones y recomendaciones.* ➘ *Se otorga importancia al estudio en profundidad de tres centros educativos (Pág.30) pero no se especifica en qué consiste ni se aprecia de forma diferencia qué resultados produjo para la evaluación.* ➘ *La definición operacional de los criterios, parámetros e indicadores (Anexo) con escalas de valoración no se aprecia de manera clara y sistemática en la exposición de hallazgos. Es probable que hayan tenido un mayor uso y reflejo en los estudios de caso, si bien no se ha tenido acceso a estos al no anexarse al informe.* • *Entre las limitaciones de la evaluación no se referencia la disponibilidad de información o la precisión del alcance que sí aparecen señalados como Lecciones aprendidas, en el sentido de que no contribuían a alcanzar los propósitos de la evaluación. Se intuye por el contrario, aunque no se explicita, que la falta de información disponible, y el escaso soporte documental de los proyectos, han sido limitaciones importantes que se han enfrentado, a no ser que se abordasen en los estudios de caso.* • *Aprovechamiento desigual de la información recabada, especialmente la referida a los diversos cuestionarios.* • *El enjuiciamiento del proyecto a partir de los criterios definidos (eficacia, pertinencia, sostenibilidad, calidad.) queda diluido en otras lecturas del proyecto (especialmente en los dos niveles de análisis y los diferentes ámbitos de indagación establecidos), encontrando diferentes acepciones de los criterios (una las que define explícitamente y otra la que aplica, especialmente en el caso de pertinencia). Incluso los ámbitos de indagación establecidos en el diseño de metodológico no se explicitan en la exposición de las recomendaciones, lo que no facilita el examen de la trazabilidad de los resultados, restando utilidad al importante trabajo realizado de tratamiento sistemático de la información recabada, especialmente la de naturaleza cualitativa.* • *La escala de valoración de resultados del cuestionario a alumnos esta balanceada con una escala Likert de tres categorías, siendo la central está "parcialmente de acuerdo" (para los estudiantes de ciclo diversificado) y "DE acuerdo" (para los estudiantes de I y II Ciclo)*

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	Se trata de una evaluación mixta, pues combina una consultoría externa y la colaboración de MIDEPLAN y la GIZ -que podrían considerarse ajenos a la implementación del proyecto- con la participación y coordinación del proceso evaluativo por parte del MEP, que en última instancia es la institución responsable del proyecto evaluado. No obstante cabe apuntar que en el informe se califica como evaluación externa, por lo que se otorga esta calificación.
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	**5**	
		Evaluación externa (100% externo a instituciones evaluadas)	10	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y análisis de datos y presentación de resultados.	Ninguna / No aplica	0	Las fuentes elegidas contemplan una pluralidad de perspectivas, implicando a los principales stakeholders identificados en dinámicas participativas, en técnicas cuantitativas y, esencialmente, en técnicas cualitativas que predominan en los estudios de casos.
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	5	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos)	**10**	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
2.1	**¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos?** **Fases: planificación e implementación.**	Ninguna / No aplica	0	En cuanto al diseño y planificación de la evaluación, según se desprende del Informe, estuvieron implicados los agentes representados en el CAEI, con los que se realizaron dos talleres iniciales, uno de ellos en mayo de 2013 en el que, entre otros aspectos, se especificaron las preguntas de evaluación, y otro en junio de 2013 que sirvió para delimitar el problema que enfrentaba el proyecto. En la Pág.31 del informe se apunta que: "Además, se realizaron tres talleres con la Comisión de Apoyo a la Educación Inclusiva, los dos talleres iniciales permitieron delimitar el proceso evaluativo y establecer, entre otros elementos las principales preguntas de la evaluación (Ver Anexo 7)". *No hay información específica en el informe para determinar con precisión los grupos meta representados en la implementación, pues no se especifica en detalle quienes conforman la CAEI, ni cuáles son los grupos meta predefinidos. La referencia más directa en este sentido es la que se realiza en la pág. 19: "El proceso evaluativo de esta intervención permitió la participación de diversos actores vinculados a centros educativos, Nivel Central del MEP, Académicos universitarios, Instancias ejecutoras de la Evaluación, las cuales se detallan en la Tabla 2 quienes expresaron sus expectativas de participación en la evaluación a partir de sus intereses mayores. (Ver Anexo 2 en relación con el contenido específico de sus expectativas).* *Si se recurre a las técnicas aplicadas y a los instrumentos que se anexan, se pueden destacar también los siguientes grupos de agentes implicados:* ▪ Directores de centros educativos participantes en el proyecto ▪ Personal Técnico Administrativo y Administrativo de centros educativos participantes en el proyecto ▪ Padres de familia ▪ Equipos Institucionales ▪ Instancias de apoyo pedagógico ▪ Comisión de Apoyo a la Educación Inclusiva ▪ Docentes ▪ Estudiantes de Primaria, ciclo diversificado, I y II ciclo *Igualmente se apunta que "no se incluyen miembros de la comunidad local, pues no identificaron de parte de las instituciones educativas instituciones vinculadas directamente con el desarrollo del proyecto" Pág.30.*
		Ninguna / No aplica	0	
		Participan escasa o desigualmente en planificación o implementación	5	
		Participan activamente en la planificación y en la implementación	**10**	

2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica Las preguntas ni son claras ni adaptadas Las preguntas o son claras o adaptadas **Las preguntas son claras y adaptadas**	0 1 5 **10**	*Las preguntas están adaptadas al contexto y pueden considerarse claras, si bien se estima que de las 4 preguntas principales que motivan la evaluación dos de ellas se centran en elementos que pueden considerarse propios de una evaluación (esencialmente en la relativa a lecciones aprendidas y en menor medida la relativa a sistematización). Aunque no es habitual que formen parte de una evaluación este tipo de preguntas, se considera que el contexto evaluativo las demandaba y por tanto están adaptadas, intuyéndose que existía escasa información acerca de qué había ocurrido con la aplicación de este proyecto.* *Otro aspecto positivo de las preguntas de evaluación es que sufrieron un proceso de ajuste con los stakeholders, al menos con la CAEI, con la celebración de dos talleres iniciales con esta finalidad.*
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica** Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción) Los TdR presentan una formulación clara (estructura) o comprensible (redacción) Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	**0** 1 5 10	No se dispone de los TdR entre la documentación de base de esta metaevaluación. No obstante, de la lectura de las limitaciones de la evaluación que se incorpora en el informe así como de las lecciones aprendidas, se deriva que el equipo evaluador percibió problemas relevantes en cuanto a la delimitación del alcance de la evaluación y la necesidad de haber contemplado la evaluabilidad previa, para evitar un encargo falto de foco o evitar frustrar expectativas de los stakeholders.
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica No hubo un taller de inicio con stakeholders para facilitar su participación No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.) **Hubo un taller de inicio con stakeholders**	0 1 5 **10**	Según se indica en el Informe fue necesario reconstruir con la CAEI la problemática que atiende el proyecto y establecer sus antecedentes, identificar las necesidades que atendería la intervención (proyecto), la relación con sus objetivos y establecer una línea de base del proyecto (Pag.19). De esta forma se celebró un taller inicial (Taller CAEI 5 de junio de 2013) cuyos resultados se exponen en pág. 23 y Anexo 3. Al menos la CAEI fue por tanto involucrada en este Taller.
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica No son accesibles para el público (ni Informe ni resumen) Son accesibles para el público (Resumen o extracto) **Son accesibles para el público (Informe y resumen)**	0 1 5 **10**	Se entiende por accesibilidad su disposición a través de internet. En este sentido, en Internet es posible acceder al informe y al resumen que este contiene. Enlace web a Informe y resumen: https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/4dd6dda4-f69f-47e7-a009-0e5e860ba3d6/informe%20de%20evaluacion%20de%20Proyecto%20Educ%20Inclusiva.pdf?guest=true http://docplayer.es/12357423-Proyecto-apoyo-a-la-gestion-pedagogica-a-centros-educativos-de-calidad-con-orientacion-inclusiva.html
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**			
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica No se describe el objetivo de la evaluación (Ni TdR ni Informe) **Se describe el objetivo de la evaluación de manera imprecisa** Se describe de manera precisa el objetivo	0 1 **5** 10	Se describen en las pág. 8 (Resumen ejecutivo) y 19 (Informe final) 1. Establecer los logros de los objetivos del Proyecto "Apoyo a la Gestión Pedagógica de centros educativos de calidad con orientación inclusiva". 2. Valorar la sostenibilidad de las prácticas impulsadas y los resultados alcanzados por el proyecto. 3. Identificar lecciones aprendidas que impulsen la gestión de la transformación de los centros educativos hacia la educación inclusiva. 4. Sistematizar la experiencia educativa del proyecto como insumo de divulgación a otras audiencias. 5. Favorecer un aprendizaje crítico de los lineamientos orientados hacia la evaluación impulsados por MIDEPLAN. Aunque si se aprecia que se describe el objetivo de la evaluación, debe apuntarse que los evaluadores señalan que dichos objetivos hubiesen requerido un análisis de evaluabilidad previo así como una mayor acotación del alcance evaluativo, algo que presentan como una limitación del proceso evaluativo y como una lección aprendida. En este sentido resulta relevante la siguiente cita *"Sin embargo, en el marco del desarrollo de la evaluación, particularmente en la delimitación de la teoría de la intervención, se reconoce que los objetivos de la evaluación propuestos no son consistentes con dicha propuesta,*

				por eso es necesario generar un espacio para articular los propósitos de intervención prescrita con la expectativa del foco de la evaluación. Esta articulación implicó reconstruir con la CAEI la problemática que atiende el proyecto y establecer sus antecedentes, identificar las necesidades que atendería la intervención (proyecto), la relación con sus objetivos y establecer una línea de base del proyecto". Por ello se otorga la calificación intermedia, vista la dificultad que entrañó su definición para los evaluadores.
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica No se describe el objeto de evaluación **Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes** Se describe con precisión (al menos los ítems contemplados)	0 1 **5** 10	De los ítems que explícitamente incorpora esta pregunta (grupos interesados, dimensión regional, etc.) en el Informe no se contempla el volumen financiero (sólo hay una única referencia al volumen financiero del año 2011 de una de las entidades financiadoras). El resto de ítems si son contemplados y se puede afirmar que presenta una descripción precisa. Dicho esto, se considera que la descripción del objeto es mejorable para facilitar al lector una comprensión completa del proyecto, su filosofía, actividades y conceptos propios. Haber contado con una definición previa de conceptos o un mayor detalle de los mismos en la descripción del objeto, hubiese facilitado la comprensión. Aunque alguno de estos conceptos se van explicando a medida que discurre la lectura del informe. Conceptos como el amigo crítico, en qué consiste realmente el índice de calidad o incluso el rol de determinadas entidades que se van nombrando, son algunos ejemplos. Ya avanzado el informe el lector encuentra respuestas a estas dudas iniciales que se plantean.
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica No se realizó un análisis de stakeholders Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.) **Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)**	0 1 5 **10**	Se realizó el análisis de los grupos de interés. Se analizaron especialmente las expectativas de participación de los agentes en la evaluación, lo que se desarrolla expresamente en Pág. 19 y 20 (Tabla 2) y en el Anexo 2 (Pág. 74 a 77). En la pág. 22 se definen los actores involucrados en el problema. No se formula explícitamente el detalle de roles y funciones de dichos actores, si bien se van presentando en algunos casos a lo largo del informe. Es decir, la exposición del análisis de los actores está más centrada en el proceso evaluativo que en el objeto evaluado. Es por ello que se le ha otorgado una valoración intermedia en este ítem. No obstante, el análisis cualitativo presenta en muchos casos la sistematización de los aportes recabados por actores según roles generales adoptados en el proyecto, motivo por el cual se le otorga la valoración más alta en este ítem, en tanto que sí utiliza la perspectiva de agentes a la hora de exponer muchos de los hallazgos del informe.
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica No se elaboró **Se elaboró pero no se usa en la evaluación o no es concluyente** Se elaboró, se usa en la evaluación y es concluyente	0 1 **5** 10	Se elabora una teoría de la intervención y es positivo que se haya realizado con los actores implicados, especialmente con la CAEI. En esta teoría se contempla una cadena de resultados, un análisis de la problemática que enfrenta el programa y la exposición de una hipótesis. De todos estos elementos es la hipótesis en la que se aprecia una menor coherencia con lo que se espera de un elemento de esta naturaleza. En su exposición hay poca particularización según los diferentes contextos de los centros educativos implicados. Otro aspecto está relacionado con en qué medida la cadena de resultados que se expone o la propia metodología a la que se alude (CIIP) se haya erigido realmente en un eje central del proceso evaluativo, algo que se aprecia en menor medida en la exposición de hallazgos, conclusiones y recomendaciones. "Las principales limitaciones del proceso evaluativo encontraron su base fundamental en el proceso de identificación de la teoría de la intervención en virtud de que este ámbito se vincula con las posibilidades de identificación de los componentes de la planificación, dispuestos por MIDEPLAN, en el Ministerio de Educación Pública, en la diversidad de los intereses evaluativos manifiestos por la CAEI y, por tanto, la complejidad que representa en el ámbito educativo".
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica No se formularon **Se formularon pero no son concluyentes** Se formularon y son concluyentes	0 1 **5** 10	Se formularon hipótesis que aparecen descritas en la reconstrucción de la teoría de la intervención. Se tratan de hipótesis de acción más que específicamente de hipótesis de impacto las que se explicitan en la evaluación. Explícitamente se expone la siguiente "… como hipótesis de acción se consideró que "La implementación del Índice de Inclusión, como eje del proyecto, permite a cada centro educativo la generación de prácticas, el establecimiento de políticas y la transformación de la cultura para alcanzar la educación inclusiva". Se considera que los logros y la sostenibilidad alcanzada por los Centros educativos a partir del desarrollo del proyecto pueden analizarse desde el desarrollo de prácticas, políticas que generen transformaciones en la cultura institucional".
	¿Se diferencia entre los diferentes niveles de impacto?	Ninguna / No aplica **No se diferencia entre niveles de impacto**	0 **1**	Se contemplan diferentes niveles de impacto pero no se sistematiza con claridad el uso de diferentes niveles, aunque en la lectura está implícito. El Informe se centra más en resultados y efectos directos.

		Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)	5	"Por eso, para lograr establecer los alcances de esta evaluación se consideró sustantivo tomar como marco de referencia conceptual las indicaciones del Índice de Inclusión en cuanto a sus fases y su implementación e indicadores para poder reconocer los logros en los centros educativos y las opciones de sostenibilidad (Ver Anexo 4). La ausencia de procesos diagnósticos previos a la intervención no permite reconocer impactos y efectos que puedan ser atribuibles de manera directa a la intervención, si bien, si es posible analizar logros y rutas de aprendizaje construidos por cada centro educativo". Pág.27 No se puede hablar por tanto de niveles de impacto en sentido estricto, aunque si establece diferentes niveles de logros de los centros educativos participantes.
		Se diferencia con claridad los diferentes niveles de impacto	10	
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica	0	Se describe el contexto de la evaluación y se relaciona con la teoría, pero no es algo central……
		No se describe el contexto de la evaluación	1	
		Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención	**5**	
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	10	
3.2	**Diseño y métodos**			
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (*Los TdR no explicitan los intereses de la evaluación*)	0	No se dispone de los TdR para tomar como referente cierto cuáles fueron los intereses (explícitos) del proceso evaluativo. Se trata de una evaluación mediante estudios de caso, lo cual se considera que se corresponde con los intereses de la evaluación (entre los que destaca la sistematización y las lecciones aprendidas). En la pág. 27 se apunta que: "El proceso de evaluación del proyecto "Apoyo a la gestión pedagógica de Centros educativos de calidad con orientación inclusiva" consideró que los contextos de intervención educativa requieren diseños evaluativos que focalicen la dinámica y la complejidad de estos ámbitos, por ello, requiere la articulación de métodos cuantitativos con métodos cualitativos pero, ante todo, la participación activa de los sujetos que construyen el contexto educativo. Por lo anterior, estableció como marco el diseño evaluativo el Modelo Evaluativo: Modelo CIPP". Asimismo, en las lecciones aprendidas (pág. 16) se indica que: "Es necesario precisar el alcance de las evaluaciones públicas, con el fin de contar con un marco de orientaciones de naturaleza ética y metodológica, que permitan superar problemas tales como: la no existencia de objetos de evaluación claramente delimitados, el acceso y calidad de la información, el conflicto de intereses". Se indica así que los objetivos de la evaluación necesitaban una mayor clarificación. También, se indica que "Es fundamental realizar estudios de evaluabilidad previos para determinar el contexto de la evaluación y así determinar con mayor precisión el diseño para que sea viable técnica y económicamente" (Pág. 16). Estas apreciaciones ponen de relieve las dificultades que se enfrentaron para casar expectativas y diseño de la evaluación. Entre las principales limitaciones del proceso evaluativo se señaló precisamente "la diversidad de los intereses evaluativos manifiestos por la CAEI y, por tanto, la complejidad que representa el ámbito educativo" (Pág.35).
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	5	
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	**10**	
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica	0	Además de la referencia al Modelo CIIP (Stfufellbeam), expuesta en la observación anterior, se detalla y justifica la elección de métodos. Concretamente se apunta (Pág. 27): "Se tomó el Estudio de Caso como estrategia metodológica de indagación, recolección y análisis de la información y dadas las condiciones de desarrollo del proyecto ..Hernández (2001) y Stake (2007) .... Permitió identificar a cada institución [Centro Educativo]- como una unidad de análisis para el proceso evaluativo". (Los corchetes son propios). "Asimismo, para potenciar las ventajas metodológicas se desarrollaron estudios en profundidad en tres instituciones -Centros educativos-". No obstante no especifica en qué consisten los estudios de caso. Si bien se considera que la opción metodológica global de optar por estudios de casos como eje de la evaluación está justificada, no puede decirse que la articulación que adoptan los estudios caso esté suficientemente detallada, a lo que no ayuda que no se anexen al Informe los 10 estudios de caso.
		No se explicitan ni justifican el diseño y los métodos aplicados	1	
		Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados	**5**	
		Se explicitan y justifican suficientemente el diseño y los métodos aplicados	10	
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron	Ninguna / No aplica	0	
		No se consideró lo contrafactual	**1**	

	grupos de comparación con posterioridad?)	Se consideró tangencialmente lo contrafactual o no se especifica suficientemente cómo se consideró Se consideró lo contrafactual y se especificó cómo	5 10	Según se indica en la página 27: "La ausencia de procesos diagnósticos previos a la intervención no permite reconocer impactos y efectos que puedan ser atribuibles de manera directa a la intervención, si bien, si es posible analizar logros y rutas de aprendizaje construidos por cada centro educativo". Es decir, hay explícitamente una justificación de porque no se utiliza lo contrafactual, si bien tampoco se optó por comparar la situación de estos centros con otros no adscritos al programa, al menos en los resultados que perciben los estudiantes para inferir si sus percepciones en las escalas tipo Likert aplicadas son significativamente diferentes. Puede decirse que se contempló lo contrafactual pero se descartó su uso, siendo al menos esta opción metodológicamente más solvente que ni si quiera contemplar esta posibilidad. No obstante, se considera que existían opciones para haber contemplado lo contrafactual es probable que el foco de la evaluación en la sistematización haya decantado al equipo evaluador a explorar esas alternativas.
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica Los métodos aplicados no cubren aspectos cuanti ni cuali Los métodos aplicados cubren aspectos cuanti o cuali **Los métodos aplicados cubren aspectos cuanti y cuali**	0 1 5 **10**	Si se aplican y se cubren aspectos cuantitativos y cualitativos. Está detallado en la siguiente pregunta. Hay que decir que si bien la estrategia combina cuanti y cuali, es lo cuali lo preminente en el proceso evaluativo, incorporándose lo cuantitativo como algo complementario. No teniendo esto que ser ni positivo ni negativo, si resulta limitado el uso y explotación que se hace de la información cuantitativa, según se desprende de los contenidos de los instrumentos, especialmente cuestionarios, que aportan ítems que no están explotados en el informe. Es difícil valorar esto en su justa medida ya que descansando la estrategia metodológica en los estudios de caso, estos no se incorporan como anexos si quiera al informe. No queda claro tampoco a que se refiere la descripción de la metodología cuando dice que hace 3 Estudios en profundidad de tres centros, pues se indica que los estudios casos abordaron a los 10 centros que participan en el programa, intuyéndose que pueden referirse a estudios de casos con un mayor nivel de detalle.
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) Entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿De qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica Se aplicó sólo un método Se aplicaron dos métodos **Se aplicaron una pluralidad de métodos (más de dos)**	0 1 5 **10**	***Evaluación mediante Estudio de casos y Métodos Mixtos:*** *- Talleres con la Comisión de Apoyo a la Educación Inclusiva (3)* *- Entrevista grupal con la CAEI* *- Entrevistas: 3 (Directores de centro, Personal Técnico-Administrativo, Equipos institucionales, instancias de apoyo pedagógico -UNED, Equipo Regional Itinerante DRE Heredia, CAEI)* *- **Estudio de caso**: (10) En cada uno se aplicaron las siguientes técnicas (Ver tabla 5 Pág. 32)* *➔Visita a los centros educativos (10, 1 por caso)* *➔Entrevista grupal: 28 en total= 8 con familias (38 participantes); 10 con personal técnico administrativo (78 participantes); y 10 con Equipo institucional de Centros educativos (71 participantes)* *➔Entrevista personal: 10 en total (Director. Centro),* *➔ Taller: 18 en total = 10 con docentes (100 participantes); 8 con estudiantes (163 participantes);* *➔ Encuesta: 3 en total aplicadas en los 10 casos: 1 a docentes (185 en total), 1 a estudiantes (611 en total) y 1 a familias (286 en total)* *- Estudios en profundidad de Centros educativos: 3*
3.3	**Procedimiento de selección**			
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección **Se explicitan en algunos casos (técnicas)** Se explicitan los procedimientos y criterios de selección	0 1 **5** 10	*En general puede decirse que las elecciones metodológicas relevantes encuentran alguna justificación explicita de los criterios de selección, tanto de los centros participantes, como de aquellos que son objeto de un estudio en profundidad. También con los stakeholders considerados (los que se derivaron de los talleres con la CAIE)* *No obstante, es en las encuestas realizadas donde estos procedimientos y criterios son menos explícitos.*
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? **Se entiende que aplica esta pregunta para las Encuestas**	Ninguna / No aplica La muestra no es adecuada **La muestra no es adecuada o el error no es aceptable** La muestra sí es adecuada y el error aceptable	0 1 **5** 10	El informe tiene una orientación eminentemente cualitativa, siendo los elementos cuantitativos complementarios. Esta orientación deriva en que preste escasa atención a aspectos como los que refiere esta pregunta, pues aunque se aplicaron cuestionarios, no se indica en ningún momento su representatividad sobre la población a la que refiere cada una de las muestras contempladas en cada centro educativo. La información más detalla al respecto es la que se ofrece en la página 32, resultando insuficiente. En ella se aprecian diferencias notables en la participación por centro educativo y perfil participante en encuestas (desde 7 a 34 docentes, desde 27 a 132 estudiantes; desde 5 a 110 familias), sin que medie justificación expresa del equipo evaluador. No

				obstante, se realizó un esfuerzo importante, accediendo mediante encuesta a 611 estudiantes, 286 familias, y 185 docentes. Es por ello que se otorga una puntuación intermedia en este ítem, juzgando más la escasa atención prestada a exponer cuan representativas fueron las muestras de cada centro educativo y el tipo de muestreo realizado, que la bondad o adecuación de la muestra en sí.
3.4	**Análisis de datos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica	0	Se triangulan datos, y podría decirse que de manera generalizada, si bien ……
		No se triangulan datos cuanti-cuali	1	
		Se triangulan datos cuanti-cuali pero no de manera generalizada	**5**	
		Se triangulan datos cuanti-cuali de manera generalizada	10	
3.4.2	¿Se cuantifican las respuestas siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	Ninguna / No aplica	0	En ocasiones se cuantifican las respuestas pero esto convive con datos y tablas en las que esto no se realiza y se presentan datos globales por cada centro (datos absolutos) sin elaborar un promedio de los 10 casos o sin incorporar datos relativos (porcentajes sobre el promedio) que sin duda facilitan al lector tomar una idea de la incidencia de los diferentes logros en la ejecución del proyecto.
		No se cuantifican las respuestas cuando es posible	1	
		Se cuantifican las respuestas a veces cuando es posible	**5**	
		Se cuantifican las respuestas de manera generalizada cuando es posible	10	
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica	0	Explícitamente se reconoce esto como una limitación derivado de no haber hecho un ejercicio previo de evaluabilidad. En general puede decirse que en el informe se indica que han consultado los pocos soportes documentales exístetenos, si bien esto es algo que no se aprecia en la lectura del informe, en los datos que se presentan o en las fuentes de las tablas y gráficos expuestos.
		No se usaron datos del monitoreo	**1**	
		Se usaron datos del monitoreo de manera insuficiente	5	
		Se usaron datos del monitoreo de manera suficiente	10	
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica	0	El recursos a datos externos es muy limitado y casi inexistente, por ejemplo cuando se dan datos de deserción, matriculas, etc. En contexto económico de los centros, etc. Aunque este criterio se centra en el análisis de datos, los pocos datos de fuentes externas se presentan en la contextualización del proyecto.
		No se recurrió a fuentes externas	**1**	
		Se recurrió a fuentes externas de manera insuficiente	5	
		Se recurrió a fuentes externas de manera suficiente	10	
3.5	**Criterios de evaluación**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			En la página 28 se especifica que: "Se estableció como criterio marco de evaluación la Calidad de la implementación y de los bienes y servicios…" … "Se consideró en el abordaje del trabajo de campo en cada centro educativo que la recolección de información se orientaría por los criterios de evaluación: pertinencia, eficacia, equidad y sostenibilidad" ... "prescindir del criterio de eficiencia, en virtud de que no se contó con datos ...". En la Tabla 3 se especifican los criterios abordados y en la Tabla 4 se vinculan con las diferentes etapas de implementación en cada centro educativo. Los criterios se hacen explícitos en el planteamiento metodológico pero su tratamiento queda diluido a lo largo del informe. Durante la exposición de los hallazgos es habitual que se entremezcle la exposición de juicios de valor sobre dichos criterios, eficacia y pertinencia especialmente. No obstante, prima la doble visión de la gestión de la CAIE y de la ejecución de los proyectos en los centros sobre la lectura de criterios, algo que se mantiene en la exposición de conclusiones. Todo esto permite indicar que no los criterios no han sido un referente en el enjuiciamiento, apreciándose incuso algunas diferencias en los que se indica explícitamente que se considera o define cada uno de ellos con lo que en realidad se indica al efectuar conclusiones.
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	El informe define explícitamente la pertinencia como el ajuste del proyecto a las necesidades de los centros y su contexto. Si bien se apuntan algunos aspectos, el informe se centra en describir la “pertinencia” de las actividades que contempla el índice aplicado, concibiendo ésta pertinencia más bien como “utilidad” de dichas actividades, no
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	

		Lo expone de forma concluyente y comprensible	10	abordando de manera concluyente la pertinencia según las necesidades (diferenciadas o no) de los diferentes centros educativos.
	b) logró las metas definidas	Ninguna / No aplica	0	En este apartado si es concluyente, aunque centrado sobre todo en términos globales, pues no se aborda el detalle de su aplicación en cada proyecto.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	c) es eficiente	**Ninguna / No aplica**	**0**	No se aborda pues se indica que no se disponía de datos para ello. "…se considera necesario operacionalizar los criterios de evaluación y prescindir del criterio de eficiencia, en virtud de que no se contó con datos que permitieran cumplir con los aspectos expresos por MIDEPLAN". Pág.29
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	d) causó impactos (outcome e impacto) deseados y no deseados.	**Ninguna / No aplica**	**0**	Los impactos no son algo contemplado en esta evaluación, que se limita a hablar de efectos sobre los objetivos previstos. No obstante se exponen percepciones sobre efectos directos, deseados y no deseados. "Finalmente, existen una serie de resultados emergentes para la CAEI que no fueron previstos, pero que lograron ser eficaces para potenciar el desarrollo del proyecto. Entre ellos destacan el trabajo con el ERI de Heredia, específicamente, las alianzas que han logrado a raíz de la participación del proyecto. Por otra parte, su labor desde la CAEI ha permitido apoyar y retroalimentar a algunos centros educativos han construido proyectos vinculados con la educación inclusiva. También, el desarrollo del proyecto permitió la participación de los miembros de la comisión en actividades internacionales, particularmente, el en contexto centroamericano. Por último, la posibilidad de participar en un concurso organizado por la UNESCO que favoreció la obtención de recursos para producir videos, carpetas de materiales y la traída de expertos " (Pág.37)
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	e) causó impactos duraderos	Ninguna / No aplica	0	La sostenibilidad es uno de los criterios que orientación la evaluación y están presentes en su diseño metodológico, si bien reciben escasa atención en la exposición de los hallazgos (apenas una página), lo cual no permite hablar de una exposición concluyente sobre este aspecto.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	Son varias las referencias que se realizan sobre la debilidad del seguimiento y monitoreo del proyecto (pág.47): "…si bien se reconoce el uso de instrumentos como el Plan Educativo Institucional, Plan de Mejoramiento o Planes anuales en el marco de acción del MEP, estos no incorporan elementos como metas e indicadores para asegurar su seguimiento y nivel de logro". PÁG. 47 "…No obstante, las nociones y estructuras de planificación utilizadas no permiten establecer el grado de avance de la ejecución de lo programado, ni establecer cuáles acciones no fueron implementadas, tampoco el porcentaje de ejecución de las metas de los instrumentos de planificación". PÁG. 47 El hecho de que se reconozca que ningún centro se encuentra en la implementación del Índice de inclusión en su etapa de seguimiento y evaluación (Pág.48) puede estar en la base de esta deficiencia observada. Entre las conclusiones y recomendaciones se indica alguna referencia a la necesidad de reforzar los procesos de planificación para lograr indicadores o metas como elementos de seguimiento y garantías para establecer la eficacia de las acciones desarrolladas (Pág.61) No obstante, llama la atención que no se muestre como una limitación de la evaluación ni recibe un tratamiento diferenciado en el informe esta debilidad de los sistemas de información disponibles.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	

		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	**5**	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	Puede decirse que el informe responde directamente a 3 de las 4 preguntas generales, dejando la relativa a la sistematización de la experiencia como aquella que aborda en menor medida. Esto puede estar asociado al hecho de no incorporar los estudios de caso de los 10 centros en el informe pues cabe esperar que en ellos si se haya abordado este aspecto en mayor medida.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	El informe destaca en mayor medida aspectos positivos de la experiencia aunque también contempla aspectos negativos, puede indicarse que estos últimos reciben un tratamiento mucho más residual a lo largo del informe, centrado especialmente en la exposición escueta de las conclusiones o aprendizajes
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	Si se distingue, aunque en los hallazgos –que deben caracterizarse por la exposición descriptiva de los datos recabados- se aprecian valoraciones que se van haciendo a medida que se van presentando algunos de ellos.
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	**10**	
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	Se especifica en ocasiones y general puede decirse que escasamente. La limitación a este respecto radica en que el Informe se sustenta en los estudios de caso, que no se anexan al informe, y el tratamiento de los cuestionarios es limitado, así como el análisis de las diferencias que se aprecian según centros educativos. En descargo del equipo evaluador es preciso apuntar que señalan problemas de evaluabilidad y de acotación del alcance evaluativo. No obstante, las cuestiones relativas a la fiabilidad de los datos primarios y secundarios recabados reciben escasa atención en el informe.
		No se especifica	**1**	
		Se especifica en ocasiones	**5**	
		Se especifica en general	10	
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica	0	Sí, otra cosa es que se considere que en la exposición de hallazgos se apuntaron conclusiones y recomendaciones, algunas de las cuales se diluye o no aparece como tal en el apartado dedicado a conclusiones y recomendaciones,
		Las conclusiones no están basadas en los resultados	1	
		Alguna o algunas conclusiones no están basadas en resultados	5	
		Todas las conclusiones están basadas en los resultados	**10**	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	Sí, aunque no se facilita esta correspondencia mediantes una agrupación de las mismas
		No se llega ni de forma lógica ni comprensible	1	
		Se llega de forma lógica o comprensible pero no las dos	5	
		Se llega de forma lógica y comprensible	**10**	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	Sí, es positivo que se agrupen por actores a los que se destinan. No puede decirse que todas sean del todo concretas lo cual limita su orientación a la acción para facilitar el despliegue de acciones de mejora
		No son relevantes ni específicas	1	
		Son relevantes o específicas pero no las dos	5	

		Son relevantes y específicas	**10**	
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	Contiene un resumen, breve y centrado en conclusiones y recomendaciones y lecciones aprendidas abordando en menor medida el enfoque metodológico y la exposición sintética de hallazgos.
		El Informe no contiene un resumen	1	
		El informe contiene un resumen estructurado	5	
		El informe contiene un resumen estructurado, bien legible/comprensible	**10**	
5.	**Exhaustividad y adecuación de la documentación**			
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica	0	Los TdR no están disponibles en los anexos, tampoco forman parte de la documentación consultada en esta metaevaluación, ni se encuentran disponibles ni accesibles en la web.
		Los TdR no son accesibles en anexos	**1**	
		Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	
		Los TdR son accesibles en la lengua común de trabajo o en una traducción	10	
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	El cronograma específico de la fase de campo no se incluye en los anexos del informe. No obstante, se dedica un apartado específico al trabajo de campo y su programación, en el que se exponen para cada técnica sus principales fechas e hitos claves, en algunos casos los lugares en los que se desarrollan y las organizaciones entrevistadas/consultadas (no las personas concretas).
		No se incluyen en anexos	**1**	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	5	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	Los anexos incorporan los instrumentos que sustentan las técnicas aplicadas: "Anexo 6: Instrumentos de Recolección de Información" (guías de entrevistas individuales y grupales, talleres y cuestionarios para diferentes perfiles). **Guías de entrevistas y talleres :** ▪ Guía de Entrevista a Directores de centros educativos participantes en el proyecto ▪ Guía de Entrevista a personal Técnico Administrativo y Administrativo de centros educativos participantes en el proyecto ▪ Guía de Entrevista grupal padres de familia ▪ Guía de Entrevista Equipos Institucionales ▪ Guía de Entrevista Instancias de apoyo pedagógico ▪ Guía de Entrevista Grupal para Comisión de Apoyo a la Educación Inclusiva ▪ Guía de Taller a docentes ▪ Guía de Taller Estudiantes de Primaria **Cuestionarios** ▪ Cuestionario Dirigido a estudiantes de ciclo diversificado ▪ Cuestionario Dirigido a estudiantes de I y II ciclo ▪ Cuestionario 1 Dirigido al personal docente ▪ Cuestionario para padres de familia **Estudios de caso:** En Anexo 8 también está la Anexo 8: Guía de Informe de Estudio de Caso En cuanto a la literatura y documentación consultada se presenta una bibliografía que expone algunos de los documentos consultados (curiosamente incluye como base también los 10 estudios de caso elaborados en la evaluación y los resultados de los dos talleres), sin embargo omite muchos de los documentos que se citan a lo largo del informe (ejemplos como las citas de Stake, Kaufman …. O documentos citados como Informes de educación 2011, 2005, POA, etc.)
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	**10**	
5.4		Ninguna / No aplica	0	

	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	El informe está estructurado de manera adecuada y su lectura se facilita con gráficos y tablas aunque no presenta resúmenes de los capítulos. Estos último son especialmente pertinentes dada la multiplicidad de agentes implicados y de centros educativos estudiados. No obstante, llama la atención que no formen parte del informe o sus anexos los estudios de caso realizados, dada su relevancia en esta evaluación.
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	5	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	**10**	

4.4.3 Evaluación 03: GERMINADORA

0.1	**Nombre del proyecto / programa evaluado**	**PROYECTO INTERINSTITUCIONAL GERMINADORA DE DESARROLLO ORGANIZACIONAL EMPRESARIAL, ASOCIATIVO Y COMUNITARIO (EVALUACIÓN INTERMEDIA). EN ADELANTE "PROYECTO GERMINADORA"**
0.2	**Institución que implementa**	IMAS y el Instituto Nacional de Aprendizaje (INA), asumen la elaboración de una versión interinstitucional de un proyecto previo que ya contaba con antecedentes en la región de Brunca donde se desarrolla.. (Nota: En la pregunta 2.1 de esta ficha de registro se detallan las instituciones implicadas en su ejecución interinstitucional)
0.3	**Evaluadores**	**Equipo de Evaluación** Coordinadora, evaluadora principal: Kemly Camacho. **MIDEPLAN**: Rosaura Herrera, Ericka Valerio **GIZ**: Carol Peña, **IMAS**: Adriana Sánchez (Apoyo, enlace institucional IMAS) Centro de Evaluación Alemana (CEEVAL) de la Universidad del Sarre ofrece acompañamiento científico durante todas las fases de la Evaluación.
0.4	**Monto del proyecto/ programa**	No se especifica en el Informe, indicándose que "no existe un presupuesto para un proyecto tan estratégico como La Germinadora. …. no se puede determinar con facilidad cuál es la inversión total que se ha realizado".
0.5	**Monto para la evaluación**	*No se especifica en el informe*

0.6 Descripción del proyecto/programa evaluado (un párrafo)

Antecedentes. El proyecto evaluado, desarrollado en la Región de Brunca de Costa Rica, es fruto de diversas iniciativas iniciadas en 2005 por la Escuela Planificación y Promoción Social (EPPS), de la Universidad Nacional (UNA), para impulsar procesos de desarrollo local y diversificación de la actividad productiva de la región. En 2010 se desarrolló un "Laboratorio Organizacional de Curso (LOC)" para formar Técnicos en Desarrollo Cooperativo Comunitario (TDCC). La idea era que los graduados consolidaran una empresa de TDCC, que en una fase posterior, pudiesen ofrecer sus servicios como directores de LOT, y formar Auxiliares en proyectos de inversión (API), en la región Brunca. En 2011, IMAS y el Instituto Nacional de Aprendizaje (INA), asumen la elaboración de una versión interinstitucional del proyecto, reuniendo estos antecedentes y con el fin de impulsar una visión compartida que permitiera gestionar recursos, y por esto es formulado con el nombre; "Proyecto Interinstitucional Germinadora de Desarrollo Organizacional, Empresarial, Asociativo y Comunitario", que es el que se encuentra en proceso de evaluación.

El "Proyecto Germinadora", en su versión interinstitucional, está orientado al desarrollo empresarial y organizativo que repercuta en la dinamización de la economía local y en el fortalecimiento de la acción ciudadana para la promoción del desarrollo endógeno de los espacios locales y de sus poblaciones más vulnerables. Es una iniciativa que combina la capacitación, promoción y seguimiento, para generar ingresos a través de la iniciativa empresarial identificada masivamente en zonas en condiciones de pobreza. Encuentra su fundamento en la denominada "Metodología de Capacitación Masiva (MCM)" que procura la estimulación de la conciencia organizativa de la población participante mediante el desarrollo de las capacidades de organización empresarial y ciudadana articuladas en los denominados "Laboratorios Organizacionales –LOT-". Se dirige a población adulta, especialmente de familias cuyos jefes y jefas de hogar perdieron la oportunidad de la educación formal, y se ven desplazados por las altas exigencias del mercado y los avances tecnológicos. Por ello, busca una alfabetización empresarial y técnica que permita a las personas emprendedoras, ubicar y desarrollar proyectos en nichos de mercado que estimulen los encadenamientos locales, enriquezca la actividad económica local, genere ingresos básicos a las familias y, al mismo tiempo, desarrollen capacidades empresariales.

0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL (máximo media página)

Evaluaciones piloto. Aprender-haciendo. El informe indica que "Este proceso ha sido particular ya que está acompañado de un objetivo de fortalecimiento de capacidades en evaluación para las personas que han integrado el equipo de trabajo y que son representantes del IMAS y del Programa FOCEVAL (MIDEPLAN y la GIZ), bajo la modalidad "aprender haciendo"". (Pág.6 y 39)
"Dada la naturaleza de esta evaluación … se realiza por medio de un evaluador externo, dos funcionarias de la Unidad de Acciones Estratégicas de MIDEPLAN, una funcionaria de la GIZ-FOCEVAL y una contraparte técnica institucional del IMAS, con el objetivo de aplicar la metodología de aprender haciendo de modo que los funcionarios desarrollen capacidades en evaluación. (Pág.39)

Aspectos positivos:

- "Se realizó un muy buen proceso aprender-haciendo sobre todo en la parte del diseño evaluativo" . (Pág. 134)
- "Al aplicar una metodología de aprender-haciendo se fomenta el desarrollo de capacidades en los representantes institucionales de MIDEPLAN, e IMAS de tal forma que puedan replicar dicha experiencia en futuras evaluaciones de proyectos estratégicos, y en la mejora de su desempeño en la ejecución de sus competencias institucionales". (Pág. 45)

Limitaciones o condicionantes:

- "El proceso de aprender-haciendo provoca que el proceso evaluativo sea más lento y se deba someter el avance … de acuerdo con las capacidades formadas y el nivel de asimilación de contenidos que demuestre parte del equipo evaluador." (Pág. 47)
- "El equipo evaluador no está únicamente dedicado a la evaluación del Proyecto germinadora … podría limitar su disponibilidad y flexibilidad para asumir más allá de medio tiempo laboral previsto para invertir en la evaluación. (Pág.47)
- "Debido a la extensión de esta evaluación se fue complicando la disponibilidad de tiempo de todo el equipo para que este continuara involucrado el medio tiempo de cada integrante como se había propuesto originalmente. Esto produjo que el proceso se alargara más de lo esperado". (Pág. 134)

Reflexiones finales sobre este proceso evaluativo.

- La reconstrucción de la teoría del programa del proyecto se percibe como una labor muy importante que consume muchos recursos y tiempo y que fue posible "por el aporte de FOCEVAL y al contexto de aprendizaje y de pilotaje en el que se enmarca esta experiencia. Sin embargo, los recursos dedicados a este proceso evaluativo fueron insuficientes, se ha requerido de mucho tiempo extraordinario aportado por la evaluadora principal y la GIZ."

En términos de las capacidades de evaluación, además de la necesidad de formular mejores proyectos y fortalecer el seguimiento de los mismos es indispensable para los equipos de evaluación fortalecer las capacidades de investigación y garantizar el manejo de las herramientas cuantitativas y cualitativas y su combinación. Sin una fortaleza investigativa se reduce las posibilidades de una investigación rigurosa que alimente una buena toma de decisiones y el mejoramiento de las acciones.

0.8 Cualificación cuantitativa (con comentarios) según criterios

78,8 [Sobre 100]	**10**	*Imparcialidad e independencia:* [Sobre 10]	La evaluación presenta la máxima puntuación en los dos ítems que desagregan este criterio, pues es de tipo externo (IMAS, con responsabilidad en el proyecto, funge como enlace institucional con el equipo evaluador) y utilizó fuentes que permiten capturar una pluralidad de perspectivas en el levantamiento y análisis de datos así como en la presentación de resultados.
	7.5	*Equidad, participación y transparencia:* [Sobre 10]	La evaluación obtiene una buena calificación en este criterio pues participan activamente en su planificación e implementación grupos interesados (especialmente la Comisión Interinstitucional) a través de talleres para la reconstrucción de la teoría de la intervención y se conforman círculos hermenéuticos para valorar los hallazgos con stakeholders. Aunque no se especifica que se llevasen a cabo talleres de inicio con grupos meta y stakeholders, se contemplaron los talleres citados para el diseño de la teoría de la intervención. En cuanto a las preguntas de evaluación son claras y adaptadas al contexto. No se dispone de TdR por lo que no se puede juzgar su claridad y comprensión. Y el informe está accesible en internet, no así el resumen ejecutivo que no forma parte del informe.
	37,5	*Credibilidad y exactitud:* [Sobre 50]	Destaca la evaluación en la delimitación y descripción clara de sus metas, objeto y modelo de cambio (8,7), presentando un completo análisis de los grupos de interés, su elaboración de la teoría de la intervención, diferenciando niveles de impacto, y presenta una descripción exhaustiva del contexto evaluativo. Presenta una buena valoración en términos de credibilidad y exactitud en el diseño y los métodos aplicados (8,0), pues su diseño se corresponde con los intereses de la evaluación, se aplicaron diferentes métodos que cubren aspectos cuantitativos y cualitativos. En cuanto a la consideración de lo contrafactual, lo contempló aunque descartó un diseño cuasiexperimental por problemas con la información y construyó en su defecto una comparación de poblaciones API y no API mediante encuestación. En cuanto a procedimientos de selección presenta una valoración intermedia (5), pues sólo explicita en algunos casos los criterios aplicados (encuesta y sondeo) pero no en el resto de técnicas cualitativas; y en la muestra de la encuestas siendo representativa se contempla un error muy elevado (7%) y establece un sondeo a conveniencia con un número limitado de No API. En cuanto al análisis de datos obtiene una calificación suficiente (6,25), pues triangula la información que se recaba a través de diferentes técnicas, se cuantifican a veces las respuestas que se exponen, hace un uso limitado de la información de monitoreo y recurre en algunos casos a fuentes externas para la comparación. Finalmente, presenta credibilidad y exactitud en el tratamiento y enjuiciamiento de los criterios de evaluación (9) que le son de aplicación (a excepción de los relativos a impactos y sostenibilidad).
	16,5	*Calidad del informe y utilidad:* [Sobre 20]	La evaluación de la calidad y utilidad presenta resultados notables (8,71), pues el informe da respuesta adecuada a todas las preguntas de evaluación, describe de forma balanceada sus fortalezas y debilidades, diferencias valoraciones descriptivas y valorativas y existe trazabilidad entre hallazgos, conclusiones y recomendaciones. Las recomendaciones que emite son relevantes y su grado de concreción es mejorable así como su orientación a la acción (priorización, identificación de destinatarios). Como debilidad se destaca la ausencia de un resumen ejecutivo en el propio informe, lo que merma notablemente sus registros en utilidad y facilidad de uso.
	6,5	*Exhaustividad y adecuación de la documentación:* [Sobre 10]	En el criterio de exhaustividad y adecuación de la documentación la evaluación recibe una calificación intermedia (5,5), en tanto que no presenta en anexos los TdR y en estos sólo incluye algunos de los documentos y datos que especifican el trabajo de campo realizado, aunque si anexa los instrumentos que aplica. Presenta una estructura adecuada y su lectura está facilitada por gráficos y tablas aunque no con resúmenes de los capítulos.

0.9 Resumen de fortalezas de la evaluación.

- *Amplio y sistemático trabajo de campo cualitativo, abarcando un amplio espectro de stakeholders.*
- *Sistematicidad del proceso de enjuiciamiento y análisis evaluativo, a partir del establecimiento de categorías de análisis y variables/indicadores en el diseño, que durante la implementación ayudan a vertebrar la exposición de hallazgos, facilitando el examen de la trazabilidad de hallazgos, conclusiones y recomendaciones por pregunta evaluativa.*
- *La sistematicidad del tratamiento de la información cualitativa recabada.*

0.10 Resumen de debilidades de la evaluación.

- *El amplio margen de error en la muestra de API (+-7%) junto a la realización de un sondeo con una muestra a conveniencia, hacen que algunas de las conclusiones que se emiten con base en la comparación de ambas, deban tomarse con cautela*
- *La conformación de la comparación API con No API a partir de sondeo a conveniencia deriva en una muestra de 41 personas que, aunque se explicitan sus fines, requiere cautela al establecer comparaciones.*
- *Al ser uno de los intereses de la evaluación conocer la replicabilidad del proyecto en otras regiones y cantones, el diseño evaluativo no presenta la validez externa necesaria, si bien justifica la imposibilidad de haberlo contemplado.*
- *En la exposición del objeto evaluativo se introducen conceptos que hubiesen requerido una explicación previa (Laboratorios organizaciones –LOT, ¿en qué se diferencia de los LOC?, curso preprofesionales, etc.- que si bien el lector encuentra a lo largo del informe o en Anexos, hace más compleja la primera lectura del documento exigiendo un esfuerzo añadido.*

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	**¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?**	Ninguna / No aplica Evaluación interna (100% trabaja en instituciones evaluadas) Evaluación mixta (externos y trabajan en instituciones evaluadas) **Evaluación externa (100% externo a instituciones evaluadas)**	0 1 5 **10**	Según se apunta en el informe "Se trata de una evaluación intermedia, externa, realizada en 2013 por un equipo evaluador con participación de representantes del Ministerio de Planificación (MIDEPLAN) y de la Cooperación alemana para el desarrollo (GIZ) en el marco del Programa de Fomento de las Capacidades en Evaluación en Centroamérica (FOCEVAL). Este equipo cuenta con el apoyo técnico del IMAS y es dirigido y coordinado por una experta evaluadora principal e independiente. Dicho equipo de evaluación tiene además el apoyo del Centro de Evaluación de Alemania (CEval)". (Pág.6) En el informe se apunta expresamente que se trata de una evaluación externa: "Externa: Este proceso se realiza desde una mirada externa ya que el equipo que la desarrolla no ha tenido ninguna participación previa en el Proyecto Germinadora-" (Pág.49). Por tanto, nos ubicamos ante una evaluación externa, al estar coordinada por una evaluadora principal independiente que cuenta con la supervisión y apoyo metodológico y técnico de MIDEPLAN y la GIZ, así como con el apoyo de IMAS para el acceso a información e informantes fungiendo como enlace. En tanto que IMAS está institucionalmente vinculado al proyecto (coordinación, financiamiento, etc.) debiera haberse definido con mayor claridad su rol y funciones (facilitar el acceso a la información e informantes), y no debería formar parte del Equipo evaluador como aparece en la portada del informe, para evitar equívocos.
1.2	**¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados?** Etapas: levantamiento y análisis de datos y presentación de resultados	Ninguna / No aplica Las fuentes elegidas no permiten capturar una pluralidad de perspectivas Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas **Las fuentes elegidas permiten capturar la pluralidad de perspectivas y se aplican en todas las etapas**	0 1 5 **10**	Se tomaron fuentes primarias y secundarias, y se usan tanto en el levantamiento de datos como en el análisis de datos y en la presentación de resultados
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	***Calificación (numérico)***	**Observaciones cualitativas**
2.1	**¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos?** Fases: planificación e implementación.	Ninguna / No aplica / No hay información en los documentos consultados No hay participación Participan escasa o desigualmente en planificación o implementación **Participan activamente en la planificación y en la implementación**	0 1 5 **10**	Se indica que "El desarrollo de esta evaluación se ha basado en las expectativas que todos los actores involucrados manifestaron tener sobre la misma. Para ello se han realizado consultas y se ha construido un consenso que ha sido la base del diseño y la ejecución de la evaluación". (P.7) La participación en la planificación de la evaluación puede concebirse como la implicación en la reconstrucción de la teoría del programa. A este respecto se indica que "se han realizado entrevistas a actores claves que han permitido comprender mejor la intención del proyecto …Posteriormente, se desarrolló un taller con el equipo técnico interinstitucional con quienes se validó el flujo de procesos construido y a quienes se les solicitó realizar el ejercicio de la Predeterminación del Problema (PDP)" (Pág.21). Es esta la participación más destacada (en el diseño) pues el resto de agentes implicados los son desde una perspectiva tradicional de agentes informantes clave a través de las diversas técnicas aplicadas. En la implementación cabe destacar a nivel de participación la aplicación de los denominados en el enfoque evaluativo "Círculos hermenéuticos". **Grupos implicados.** La mayor implicación se percibe con el Comité Técnico Interinstitucional que se encuentra formado por enlaces de las instituciones implicadas en el proyecto: Universidad Nacional de Costa Rica (UNA), Instituto Mixto de Ayuda Social (IMAS), Instituto Nacional de Aprendizaje (INA), Banco Popular y de Desarrollo Comunal (BPDC), Junta de Desarrollo Regional de la Zona Sur (JUDESUR) y el Movimiento Cooperativo: Instituto de Fomento Cooperativo (INFOCOOP), Centro Nacional de Educación Cooperativa (CENECOOP R.L), Consejo Nacional de Cooperativas (CONACOOP). Además es un proyecto de interés para el Despacho del Ministro de Bienestar Social.
2.2		Ninguna / No aplica Las preguntas ni son claras ni adaptadas	0 1	Las preguntas de evaluación que se formulan en el informe son claras y están adaptadas al contexto.

	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Las preguntas o son claras o adaptadas	5	▪ ¿En qué medida la adecuación y aplicación que se ha realizado de la MCM en cinco cantones de la zona sur de Costa Rica se ajusta a las condiciones contextuales y necesidades de la población participante? ▪ ¿En qué medida el involucramiento y el compromiso de las instituciones y la coordinación entre estas han facilitado el alcance de los resultados hasta este momento? ▪ ¿En qué medida el abordaje de desarrollo de capacidades promovido por el proyecto Germinadora crea condiciones para que la población más vulnerable de la zona de influencia potencie sus posibilidades organizativas, productivas y empresariales?
		Las preguntas son claras y adaptadas	**10**	
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica**	**0**	No se dispone de documento de TdR entre la documentación de esta metaevaluación.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	No se desprende de la información contenida en el informe que se realizase un taller de inicio con la finalidad de facilitar su participación. No obstante, como se ha apuntado en el ítem 2.1, sí se realizaron talleres al inicio de la evaluación para implicar al Comité Técnico Interinstitucional en el diseño de la teoría de la intervención, es por ello que se califica el ítem con el valor intermedio **Grupos implicados**. Comité Técnico Interinstitucional
		No hubo un taller de inicio con stakeholders para facilitar su participación	1	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	**5**	
		Hubo un taller de inicio con stakeholders	10	
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica	0	Enlace web al Informe de evaluación. https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/180c917b-170d-4f31-b58c-9305d04c84e5/Evaluacion_germinadora.pdf?guest=true Nótese que el Informe no se acompaña de resumen ejecutivo, por lo que se penaliza este aspecto tanto en la pregunta específica sobre esta cuestión como en esta sobre la accesibilidad.
		No son accesibles para el público (ni Informe ni resumen)	1	
		Son accesibles para el público (Resumen o extracto)	**5**	
		Son accesibles para el público (Informe y resumen)	10	
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1 Metas, objetos y modelo de cambio				
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	El objetivo principal de esta evaluación es: "Valorar la Germinadora como proyecto estratégico para el país con el propósito de retroalimentar las siguientes acciones del mismo y generar insumos para una decisión sobre la posible utilización de su forma de abordar la situación en otras regiones". Objetivos específicos. 1. Valorar si la aplicación de la metodología de capacitación masiva que se ha realizado ha sido apropiada para la zona de influencia en la que trabaja el proyecto 2. Evaluar la contribución que las instituciones participantes y la coordinación interinstitucional ha tenido en el alcance de los resultados esperados a este punto de la ejecución del proyecto. 3. Valorar si el proyecto Germinadora genera las transformaciones esperadas en las condiciones económicas, organizativas y productivas de los grupos participantes.
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	
		Se describe el objetivo de la evaluación de manera imprecisa	5	
		Se describe de manera precisa el objetivo	**10**	
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica	0	Sí, se describen con precisión las principales dimensiones del objeto de evaluación, atendiendo a todos los ítems contemplados en esta pregunta, salvo al volumen financiero, pues según se indica la información para ello no está disponible. Es especialmente completo el apartado de contextualización y antecedentes del proyecto, más si cabe que la delimitación del propio objeto. En la delimitación del objeto, algunos de los conceptos clave del proyecto se dan por supuestos y es preciso acudir a los anexos donde se puede encontrar mayor detalle.
		No se describe el objeto de evaluación	1	
		Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes	5	
		Se describe con precisión (al menos los ítems contemplados)	**10**	
3.1.3		Ninguna / No aplica	0	

	¿Se realizó un análisis de los grupos de interés (stakeholder)?	No se realizó un análisis de stakeholders	1	Sí, se realizó el análisis de los grupos de interés y de manera bastante completa, señalando tanto los roles y principales funciones de cada grupo. Se presenta "Mapeo de actores desde las dimensiones del grado de poder e interés en el proyecto" y la identificación de los stakeholders de acuerdo con su función en el proyecto como gestores, financiadores, poblaciones finales, ejecutores o apoyo técnico (Anexo 4). También se detalla el rol y expectativas de participación de los stakeholders en l(Anexo 4).
		Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.)	5	
		Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	**10**	
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica	0	En el apartado 1.5.1 se detalla la teoría del programa de este diseño de evaluación, apoyada en tres instrumentos: a. El flujo de procesos que expresa la ejecución realizada de la Germinadora b. El análisis de la problemática que se aborda con el proyecto, basada en la aplicación del instrumento Predefinición del problema (PDP). c. La línea de intención o cuadro de resultados donde se explicitan las actividades, los productos, efectos e impactos. A esta línea de intención se le han adicionado los supuestos e insumos de los que parte el Proyecto Germinadora. Son los dos primeros instrumentos los que reciben mayor atención. Y, es la definición del cuadro de resultados el aspecto que se podría decir que recibe menor atención pues su exposición en tablas se acompaña de una breve explicación que trata de clarificar cómo se espera que unos elementos del cuadro de resultados generen aquellos otros de orden o nivel superior con los que están asociados. En este cuadro cabe apuntar que algunos de sus elementos se definen más como indicadores que como resultados La articulación de esta Teoría del Programa es particular respecto a lo que estipula el Manual de MIDEPLAN que presta atención a otros aspectos.
		No se elaboró	1	
		Se elaboró pero no se usa en la evaluación o no es concluyente	5	
		Se elaboró, se usa en la evaluación y es concluyente	**10**	
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica	0	En la explicación sobre el cuadro de resultados se enuncian (aunque no explícitamente como tales) lo que podrían ser hipótesis de impacto o de cambio. Se encuentran implícitas en la descripción de la teoría del programa. Por el contrario, presta mayor atención a la explicitación de supuestos en los que se basa el proyecto. Por todo ello, se otorga una puntuación intermedia.
		No se formularon	1	
		Se formularon pero no son concluyentes	**5**	
		Se formularon y son concluyentes	10	
	¿Se diferencia entre los diferentes niveles de impacto?	Ninguna / No aplica	0	Se diferencia entre efectos de primer nivel, de segundo nivel e impactos, apuntando que la evaluación pone su foco en productos y efectos de primer nivel
		No se diferencia entre niveles de impacto	1	
		Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)	5	
		Se diferencia con claridad los diferentes niveles de impacto	**10**	
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica	0	El contexto de la intervención se define y delimita de manera más profusa que el propio contexto evaluativo, aunque este queda bien delimitado a partir de la justificación del proceso y la definición del problema que aborda la evaluación. Se aprecia un contexto evaluativo marcado tanto por un propósito sumativo y de aprendizaje tanto para la mejora de la intervención como para valorar su extensión nacional.
		No se describe el contexto de la evaluación	1	
		Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención	5	
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	**10**	
3.2	**Diseño y métodos**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (*Los TdR no explicitan los intereses de la evaluación*)	0	No se dispone de los TdR de forma que permita formarse una opinión objetiva sobre los intereses de la misma, lo cual supone una limitación importante para emitir este juicio. Según se describe en el informe, los enfoques de evaluación seleccionados son: evaluación participativa, evaluaciones útiles y evaluaciones para la rendición de cuentas (Pág. 38). Asimismo, se indica que se ha considerado fundamental utilizar un enfoque de este tipo "debido al tipo de proyecto y al momento en que se encuentra. Factores como la no existencia de una unidad ejecutora, el no disponer de un presupuesto propio y que la ejecución esté basada en las acciones de cada institución participante hacen de Germinadora un proyecto particular que amerita una evaluación constructivista que haga emerger de la realidad aspectos no predeterminados" (Pág.52). Este diseño se corresponde con los intereses que el informe explicita para la evaluación, siendo quizás el relativo a la potencial extensión del programa el interés evaluativo que encontraría menor correspondencia en términos de validez para su extrapolación a otros contextos. No obstante, las limitaciones en cuanto a disponibilidad de
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	5	
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	**10**	

				información, apuntan a la oportunidad de reconstruir procesos, teorías y marcos de resultados, estableciendo líneas de base para futuras evaluaciones, como apunta el propio informe
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica No se explicitan ni justifican el diseño y los métodos aplicados **Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados** Se explicitan y justifican suficientemente el diseño y los métodos aplicados	0 1 **5** 10	"La evaluación estratégica del Proyecto Germinadora se fundamenta en una serie de argumentos clasificados en tres ámbitos explicativos: la mejora en la intervención (continuidad y replicabilidad), la rendición de cuentas, el desarrollo de capacidades en el tema de evaluación" (Pág.38). Este diseño realiza una combinación de métodos cuantitativos y cualitativos, y utiliza técnicas de ambos métodos con preminencia de las cualitativas, siendo su abordaje prioritario constructivista (Pág. 52). El diseño global de la evaluación está explicado suficientemente. Los métodos aplicados se explican aunque no están justificados de manera generalizada y suficiente, salvo en el caso de la tercera pregunta evaluativa, donde si hay una justificación expresa del descarte de la opción cuasiexperimental y, por tanto, de la selección del enfoque constructivista que se adopta.
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	Ninguna / No aplica No se consideró lo contrafactual **Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró** Se consideró lo contrafactual y se especificó cómo	0 1 **5** 10	Según se indica en el informe "… aunque se valoró la posibilidad de trabajar esta interrogante (-se refiere a la tercera interrogante evaluativa relativa al desarrollo de capacidades) con un diseño cuasi-experimental se consideró que las condiciones en las que se desarrolló el proyecto no lo permiten …. o se haya hecho posible determinar marcos muéstrales apropiados para poblaciones participantes y no participantes" Pág.64. La propuesta cuasiexperimental que se barajó fue contemplar una suerte de "listas de reserva" entre los API, considerando en vez de aquellas personas que se postularon para API, aquellas que participaron en el primer taller de capacitación masiva pero que decidieron no continuar. Al no disponer de la información básica de estos participantes se descartó esta opción, ya que estaba disponible sólo para 2 Laboratorios de Cantones, motivo por el cual se decidió realizar simplemente un sondeo entre ellos que permita un cierto contraste o comparación no experimental entre sus logros y los que se graduaron como API. Además, según la evaluadora principal, "otros diseños evaluativos… se ajustan de una mejor manera a la situación real y específica de este proyecto". En definitiva, puede afirmarse que se consideró lo contrafactual, aunque se descartó su aplicación cuasiexperimental, y se adoptó contemplar una comparación no experimental.
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica Los métodos aplicados no cubren aspectos cuanti ni cuali Los métodos aplicados cubren aspectos cuanti o cuali **Los métodos aplicados cubren aspectos cuanti y cuali**	0 1 5 **10**	Si se aplicó una pluralidad de métodos para el levantamiento de datos, cubriendo aspectos cuanti y cuali, si bien predominan estos últimos dado el enfoque constructivista del proceso evaluativo.
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica Se aplicó sólo un método Se aplicaron dos métodos **Se aplicaron una pluralidad de métodos (más de dos)**	0 1 5 **10**	***MÉTODOS MIXTOS:*** ***<u>Técnicas de investigación</u>*** - ***Análisis documental*** - ***Taller con el comité técnico interinstitucional para validar el proceso ejecutado*** - ***Grupo focal con Auxiliares de Proyectos de Inversión*** - ***Entrevistas semiestructuradas*** *(narrativas y con escalas tipo Likert): 62; 2 (Expertos de la Metodología de Capacitación Masiva) y 60 actores clave (9 TDCC, 20 API, 10 "Fuerzas vivas locales", 6 "Institucionalidad local, 15 "Instituciones centrales").* - ***Encuesta****: 1; a una muestra de173 API, estratificada por los 5 cantones y de los 4 años (2010, 2011, 2012, y 2013) del período de evaluación de La Germinadora.* - ***Sondeo:*** *1; a una muestra 41 no-API de los 3 cantones y los 4 años con datos disponibles.* - ***Estudios de caso****: 15; estudios de API de los 5 cantones participantes en diferentes períodos, ejecutando diversas actividades económicas, hombres y mujeres y de distintas edades.* - ***Círculos hermenéuticos****: 3; 1 con Técnicos de Desarrollo Comunitario y Cooperativo (TDCC) y 2 con Auxiliares de Proyectos de Inversión (API)*

				- ***Análisis retrospectivo**, a partir de líneas de tiempo, del proceso desarrollado en dos cantones para efectos comparativos de la aplicación metodológica.* ***Análisis de datos:*** - ***Herramientas:** SPSS (datos cuantitativos), Atlas-ti (datos cualitativos) y Tiki-toki (líneas de tiempo).* - ***Estrategia:** triangulación y retroalimentación con actores clave* - ***Otros análisis realizados*** ▪ ***Análisis y revisión de bases de datos: 2;** personas beneficiarias y propuestas de proyectos* ▪ ***Análisis retrospectivo basado en la reconstrucción de los procesos del proyecto en dos espacios:** Osa y Golfito.* ▪ ***Análisis comparativo de procesos a partir de líneas de tiempo***
3.3	**Procedimiento de selección**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección **Se explicitan en algunos casos (técnicas)** Se explicitan los procedimientos y criterios de selección	0 1 **5** 10	*Se explicitan los criterios y el procedimiento de selección en el caso de las encuestas y el sondeo, siendo especialmente detallado en las primeras, donde se exponen los criterios que determinaron la elección metodológica y los procedimientos empleados en el marco muestral definido.* *En el resto de técnicas cualitativas, que son las predominantes, no es generalizado la explicitación de su justificación en el proceso de diseño, si bien se detallan en algunas entrevistas y no en los estudios de caso. Son estos últimos los que quedan más indefinidos, dado que por definición contemplan una amalgama de otras técnicas y resulta especialmente relevante detallar como se conformaron, sus criterios, procedimientos y técnicas contempladas.*
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? **Se entiende que aplica esta pregunta para las Encuestas**	Ninguna / No aplica La muestra no es adecuada **La muestra no es adecuada o el error no es aceptable** La muestra sí es adecuada y el error aceptable	0 1 **5** 10	La muestra de la encuesta API es representativa y se muestra adecuada aunque presenta un error muy elevado y el sondeo se realiza a conveniencia con un número limitado de No API. ***ENCUESTA API:*** *Se contempla un marco muestral de la población API según el cantón en que cada uno desarrolla su proyecto productivo particular y el año en que se incorporó al proceso participando en los LOT.* ***Población total API**: 1.109 / 1.059 que iniciaron un proyecto productivo (listado oficial de API)* ***Muestra**: - Selección de distribución muestral de proporciones según año de participación (2010-2013)* *- Calculo de muestra para población finita con* *- p de 0,5 (50% de probabilidad y 50% de no probabilidad)* *- Nivel de confianza del 95% para generalizar resultados al comportamiento de la población.* *- Error máximo permisible del 7% (0,01), variando los resultados entre +7% y -7%.* *- Tamaño de la muestra: Ochenta y ocho (88) encuestas a aplicar.*__Nota__: Existe una Discrepancia entre lo indicado en texto (88) y lo que aparece en las tablas (171).* ***SONDEO NO API*** *Sólo se dispone información para 2 cantones. Se determina por conveniencia trabajar con un número de No-API que sea del 20% de los API.* *Cuadro 8* *Muestra del Proyecto Germinadora* *Cantidad de encuestas a aplicar por Año y por Cantón* *Cuadro 9* *Sondeo NO API del Proyecto Germinadora* *Cantidad de encuestas a aplicar por Cantón*

CANTIDAD DE ENTREVISTAS API					
	AÑO				
CANTON	2010	2011	2012	2013	TOTAL
Buenos Aires	14	0	12	0	26
Corredores	20	4	0	0	24
Coto Brus	19	0	12	0	31
Golfito	15	13	17	0	45
Osa	7	0	20	18	45
TOTAL	75	17	62	18	171

CANTIDAD DE ENTREVISTAS NO API	
CANTON	Todos años
Corredores	15
Golfito	28
TOTAL	43

3.4	Análisis de datos	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali Se triangulan datos cuanti-cuali pero no de manera generalizada **Se triangulan datos cuanti-cuali de manera generalizada**	0 1 5 **10**	La triangulación forma parte de la estrategia metodológica. Se triangulan los datos cuantitativos y cualitativos en la exposición de los hallazgos para cada pregunta evaluativa. No obstante cabe apuntar que en puridad de criterio, no queda claro en qué medida son las mismas categorías y variables las que se recaban a través de diferentes técnicas y no supone simplemente una contrastación de información diversa a través de diferentes técnicas.
3.4.2	¿Se cuantifican las respuestas siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	Ninguna / No aplica No se cuantifican las respuestas cuando es posible **Se cuantifican las respuestas a veces cuando es posible** Se cuantifican las respuestas de manera generalizada cuando es posible	0 1 **5** 10	El informe combina tanto la exposición de generalizaciones a partir de la información cualitativa recabada, tratada y sistematizada, con la información cuantitativa, esencialmente la obtenida a partir de las encuestas y en algún caso también se cuantifican percepciones recabadas a partir de entrevistas, pero en menor medida. Predomina la generalización sobre la exposición de cuantificaciones sobre las percepciones recabadas.
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica No se usaron datos del monitoreo **Se usaron datos del monitoreo de manera insuficiente** Se usaron datos del monitoreo de manera suficiente	0 1 **5** 10	Se usaron datos del monitoreo del proyecto que se citan en la bibliografía, entre ellos Informe General de Avance 2012, Informe de Seguimiento de Proyectos Corredores y Golfito, Fichas técnicas de proyectos productivos, Informe Final Lot Golfito Corredores 2011, Laboratorio Organizacional de Curso, Heredia (2010), Laboratorio Organizacional de Terreno, registro de participantes de cursos APIS Golfito (2012), etc. La información principal que se deriva de estos informes se contempla en los anexos del informe más que en el propio cuerpo del informe y en la exposición de hallazgos. Es por ello que la lectura de las realizaciones y resultados del proyecto se apoya más en los datos cualitativos y en las encuestas realizadas que en los propios datos de seguimiento (número y tipología de cursos realizados según las diversas modalidades y territorios; ámbitos de los proyectos empresariales; etc.)
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica No se recurrió a fuentes externas **Se recurrió a fuentes externas de manera insuficiente** Se recurrió a fuentes externas de manera suficiente	0 1 **5** 10	Se recurrió a fuentes externas especialmente en el apartado de contextualización. No obstante, por lo general no se cita la fuente de la que provienen, salvo alguna excepción (Censo de 2011). También en los hallazgos se citan estudios de MIDEPLAN y el Estado de La Nación sobre población más vulnerable de la zona sur-sur. No obstante, no es un recurso generalizado en la exposición de hallazgos y en el análisis o la triangulación, la referencia a fuentes externas que permitan la comparación o el contraste de realizaciones y logros del proyecto. Es por ello que se otorga una calificación intermedia, tanto por no referenciar las fuentes que se utilizan en la contextualización (que sí parecen suficientes para ese cometido), como por su escaso uso en el análisis de datos.
3.5	**Criterios de evaluación**	**Calificación (texto)**	***Calificación (numérico)***	**Observaciones cualitativas**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica No lo expone -1 Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente **Lo expone de forma concluyente y comprensible**	0 1 5 **10**	Una de las preguntas se dedica expresamente a la pertinencia, exponiendo hallazgos y conclusiones al respecto y recomendaciones de mejora. Otra de las preguntas hace referencia a la relevancia, si bien se define ésta más asociada a los efectos intermedios del proyecto. La mayor debilidad del enjuiciamiento de la pertinencia es que dedica poca atención a las necesidades en sí de la región/cantón y se centra más en las necesidades de los perfiles de grupos meta (edad, sexo, nivel de estudios, etc.), así como en la idoneidad de la adaptación de la metodología MCM a la zona. Las necesidades generales del territorio están expuestas en la contextualización pero no se abordan en el enjuiciamiento de la pertinencia, sino que indirectamente se indica la falta de adaptación al determinar el débil papel que jugaron los diagnósticos locales.

	b) logró las metas definidas	Ninguna / No aplica	0	En el caso de las metas se analizan tanto en la tercera interrogante evaluativo, denominada aquí relevancia como en la pregunta global. Lo expone de forma concluyente, a partir de los métodos aplicados.
		No lo expone -1	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	c) es eficiente	Ninguna / No aplica	0	Una de las preguntas se dedica expresamente a la eficiencia –entendida como la contribución de las instituciones participantes y la coordinación interadministrativa en los logros del proyecto- , exponiendo hallazgos y conclusiones al respecto y recomendaciones de mejora
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	d) causó impactos (outcome e impacto) deseados y no deseados.	**Ninguna / No aplica**	**0**	
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	Las conclusiones, sobre todo las relativas a la segunda pregunta evaluativa, hacen referencia a la necesidad de optimizar la gestión del proyecto, entre otros aspecto, sobre la recolección sistemática de información de su s realizaciones y resultados. Esta información aparece más bien dispersa por diferentes apartados del informe, por lo que hubiese sido recomendable un tratamiento que hiciese más visible esta cuestión. Cabe apuntar en descargo de la evaluación que no fue esta una pregunta expresamente formulada en el encargo evaluativo.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	La comparación no experimental realizada le permite identificar cambios, que se reconducen al proyecto, si bien se puede decir que las debilidades intrínsecas de la opción metodológica por la que se tuvo que decantar esta evaluación al no poder realizar otras, limitan un registro que pueda calificarse como concluyente a este respecto.
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	**5**	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	En el informe se responde de manera adecuada a las tres preguntas de evaluación por lo que se le asigna un 10. No obstante, cabe apuntar algunas limitaciones en el informe en cuanto a las respuestas a las preguntas en cuanto especialmente las dos primeras no contemplan comparativas o análisis de los diferentes espacios en los que se aplica el proyecto, de manera que se puedan apreciar diferentes resultados según diferentes contextos y tipos o modalidades de aplicación. Esto es algo que si se tiene en cuenta en la tercera pregunta de evaluación y a partir de la encuesta utilizada para responderla. Esto es importante de cara a ofrecer respuestas a la pregunta relativa a la replicabilidad del modelo de MCM para su extensión nacional. En este mismo sentido el limitado uso que se hace de la información de seguimiento sobre realizaciones del proyecto en los distintos cantones (información que si ha sido consultada) cuando se triangula la información recabada para exponer los hallazgos relativos a cada pregunta, más apoyados en la información cualitativa y, en algunos casos, para un mayor detalle de algunas aplicaciones concretas debe recurrirse a los anexos.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	

4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	El informe presenta de manera equilibrada puntos fuertes y débiles del programa, todo ello en función de los hallazgos que se exponen para cada pregunta evaluativa, que lógicamente en unos casos iluminan en mayor medida, puntos fuertes o débiles de este proyecto.
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	En la lectura del informe puede distinguirse cuando la exposición formula aspectos o datos más descriptivos derivados dela indagación realizada o cuando se formulan valoraciones o enjuiciamientos sobre la intervención evaluada. Dicho esto, las declaraciones valorativas no solo se encuentran circunscritas al apartado donde se exponen las conclusiones, sino que se entremezclan en la exposición de hallazgos (algo normal pues su formulación comporta en ocasiones una necesaria interpretación de datos recabados por el equipo evaluador), y también en la exposición de datos recabados, siendo esto algo que debiera evitarse en el informe.
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	**10**	
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	
		No se especifica	1	
		Se especifica en ocasiones	5	
		Se especifica en general	**10**	
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica	0	La lectura de las conclusiones muestra con claridad su anclaje en los datos recabados que han sido expuestos y en los resultados de la evaluación.
		Las conclusiones no están basadas en los resultados	1	
		Alguna o algunas conclusiones no están basadas en resultados	5	
		Todas las conclusiones están basadas en los resultados - 10	**10**	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	La lectura de las recomendaciones muestra con claridad su anclaje en las conclusiones alcanzadas. El hecho de que las recomendaciones y conclusiones estén vinculadas a cada pregunta de evaluación formulada (al igual que ocurre con la exposición de la metodología aplicada y los hallazgos), facilita esta lectura y trazabilidad del enjuiciamiento en esta evaluación.
		No se llega ni de forma lógica ni comprensible	1	
		Se llega de forma lógica o comprensible pero no las dos	5	
		Se llega de forma lógica y comprensible	**10**	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	Las conclusiones son relevantes y su nivel de especificidad es variable. A este respecto, algunas recomendaciones se formulan de manera muy concreta y específica, facilitando su aplicación a los gestores, y en otros se realiza una formulación de corte genérico que dificulta su concreción sobre todo pensando en un plan de acción, y por tanto limita la orientación a la acción que debe caracterizar a las recomendaciones. El hecho de que no haya lectura por perfiles de agentes en las recomendaciones también es otro aspecto de mejora de la concreción de las recomendaciones. Así como que se hubiese realizado una priorización de las recomendaciones que se emiten.
		No son relevantes ni específicas	1	
		Son relevantes o específicas pero no las dos	**5**	
		Son relevantes y específicas	10	
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	El informe no incorpora resumen ejecutivo.
		El Informe no contiene un resumen	**1**	
		El informe contiene un resumen estructurado	5	
		El informe contiene un resumen estructurado, bien legible/comprensible	10	
5.	**Exhaustividad y adecuación de la documentación**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica	0	Los TdR no están accesibles en los anexos ni se han aportado en la documentación de base para esta metaevaluación.
		Los TdR no son accesibles en anexos	**1**	
		Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	

		Los TdR son accesibles en la lengua común de trabajo o en una traducción	10	
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	El informe no incorpora un anexo expreso para sistematizar la información básica sobre el trabajo de campo (cronograma, fases, listados de lugares, organizaciones visitadas y personas entrevistadas). No obstante, en algunos casos si se aporta esta información o parte de ella, como en el caso de: ▪ Anexo 14. Análisis comparativo normativo-fáctico. Cita a los entrevistados ▪ Cuadro 7 y 8. Listado círculo hermenéutico. Cita a los participantes, fechas y lugares ▪ Anexos 12 y 13. Entrevistas abiertas a expertos. Cita a las personas entrevistadas. Asimismo, lo habitual es que al exponer las técnicas aplicadas en la respuesta a cada pregunta evaluativa se exponga la cuantificación del número de aplicaciones que tuvo cada técnica (número de entrevistas según perfil, número de círculos hermenéuticos según perfil, número de encuestados, etc.). También en este apartado suele exponerse las fases que contempla el trabajo de campo pero no su cronograma. Este proceder dificulta conocer de manera rápida los detalles sobre la ejecución del trabajo de campo.
		No se incluyen en anexos	1	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	**5**	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	En este apartado el Informe es completo, en tanto que en los anexos incorpora todas las fuentes consultadas y todo los instrumentos aplicados, y en algunos casos incluso suma a estos una síntesis de la explotación y sistematización de datos recabados a partir de cada instrumentos (como hace con los círculos hermenéuticos y sobre todo con las entrevistas).
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	**10**	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	Ninguna / No aplica	0	El informe presenta una estructura adecuada, salvo la excepción hecha de algunas omisiones relevantes (como la inclusión en anexos de TdR, o de un resumen ejecutivo, entre otros). Asimismo, incorpora elementos que facilitan su lectura cuando es posible, bien a través de gráficos o bien a través de tablas; no contemplando otros como resúmenes de los capítulos, salvo alguna excepción.
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	5	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	**10**	

4.4.4 Evaluación 04: PND

0.1	**Nombre del proyecto / programa evaluado**	**EVALUACIÓN DEL "DISEÑO METODOLÓGICO Y PROCESO DE ELABORACIÓN DEL PND 2011 – 2014", DE MIDEPLAN.**
0.2	**Institución que implementa (p.ej. Ministerio)**	Ministerio de Planificación Nacional y Política Económica (MIDEPLAN): responsabilidad principal de la elaboración del PND. (Art. 4º de la Ley No 5525) Implanta normas de asesoría, información y coordinación necesarias con el resto del Sistema Nacional de Planificación (SNP)[36]
0.3	**Evaluadores**	Instituto IDEA International
0.4	**Monto del proyecto/ programa**	PND:5,2 billones de colones (Pág. 20)[37]
0.5	**Monto para la evaluación**	*No se especifica en la evaluación y no se dispone de los TdR*

0.6 Descripción del proyecto/programa evaluado (un párrafo)
El objeto de esta evaluación es el Plan Nacional de Desarrollo (PND) "María Teresa Obregón Zamora" de Costa Rica 2011-2014. No obstante el interés evaluativo se centra en mejorar el diseño metodológico y el proceso de elaboración del Plan Nacional de Desarrollo (PND) 2015-2018 a partir de la experiencia del PND 2011-2014 y proponer una nueva metodología para orientar la planeación hacia resultados más estratégicos y evaluables, que permita una vinculación con marcos presupuestarios de corto-mediano plazo y que facilite una participación más amplia de diversos actores representativos sociales e institucionales. *El PND 2011-2014 presenta una estructura que busca enlazar las políticas y metas nacionales con las políticas, metas y acciones estratégicas sectoriales. MIDEPLAN tiene la responsabilidad principal de la elaboración del PND. Para ello debe implantar normas de asesoría, información y coordinación que sean necesarias con el resto del Sistema Nacional de Planificación (SNP), el cual deberá prestarle toda la cooperación técnica requerida. Así, MIDEPLAN emite una metodología para la elaboración del PND, con el propósito de que los actores, definan metas, incorporen programas, proyectos, asignen recursos y definan responsables y así ordenar los procesos de las entidades, al tiempo que permita el seguimiento y evaluación, la transparencia y la rendición de cuentas. Los avances de monitoreo y seguimiento al PND 2011-2014 muestran brechas entre los objetivos planteados y el resultado final que se elabora a partir de la interacción de los actores que intervienen en el proceso de construcción del PND. Incluso, la Contraloría General de la República ha señalado en sus informes la necesidad de mejorar los indicadores para la medición de la eficiencia y que se presentan divergencias entre indicadores, metas y estimación de recursos al comparar lo establecido en el PND 2011-2014, recomendando "mejorar la vinculación entre la planificación y la presupuestación (…) estableciendo procedimientos para la vinculación en los momentos claves de los procesos, como es la elaboración del PND".*

0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL (máximo media página)
Evaluación externa (Consultoría Internacional), apoyada por FOCEVAL I[38] El hecho de que la evaluación se realizase bajo esta modalidad de evaluación externa hace que en el informe no se preste atención al desarrollo de capacidades que se haya podido derivar del proceso, ni tampoco otras que puedan inferirse como aprendizajes derivados del proceso, fortalezas o limitaciones. Simplemente se limita a exponer que FOCEVAL, junto a MIDEPLAN, patrocinan esta evaluación y que ambas organizaciones supervisaron el desarrollo del trabajo de campo de la evaluación y sus principales instrumentos aplicados. Tampoco hay referencias a si se conformó una estructura para la gestión de la evaluación con MIDEPLAN o FOCEVAL-GIZ.

[36] Forman el SNP: a)MIDEPLAN; b) Las unidades u oficinas de planificación de los ministerios, instituciones descentralizadas y entidades públicas locales y regionales; y c) Los mecanismos de coordinación y asesoría, tales como consejos asesores, comités interinstitucionales, comisiones consultivas y otros

[37] En el informe se indica que esta estimación por ejes que contiene el PND no es consistente con la estimación presupuestal de los cuadros de Lineamientos de Políticas Sectoriales del Anexo 1

[38] Aunque los TdR de esta metaevaluación indicaban que esta evaluación fue apoyada bajo la modalidad de aprender-haciendo, se realizó bajo la modalidad de evaluación externa

0.8 Cualificación cuantitativa (con comentarios) según criterios			
73 [Sobre 100]	**10**	*Imparcialidad e independencia:* [Sobre 10]	*La evaluación obtiene la máxima puntuación en este criterio al ser de tipo externo en la que además se considera que las fuentes seleccionadas permiten capturar una pluralidad de perspectivas en sus distintas etapas.*
	6,5	*Equidad, participación y transparencia:* [Sobre 10]	*La evaluación obtiene una calificación suficiente en este criterio pues participan activamente en implementación grupos interesados –si bien como informante clave- (aunque no se aprecia la participación en el diseño). En cuanto a las preguntas de evaluación son claras y adaptadas al contexto. No se dispone de TdR por lo que no se puede juzgar su claridad y comprensión. Y tanto el informe como el resumen ejecutivo que aquél contiene están accesibles en internet.*
	32,5	*Credibilidad y exactitud:* [Sobre 50]	*La evaluación obtiene una valoración notable en este criterio (7), pues está descrita la meta/objetivo de la evaluación, describe el objeto evaluado aunque sea en algunos aspectos (no en todos) y describe exhaustiva el contexto evaluativo. Le penaliza en este sentido que no se aprecia que realizase un taller inicial con stakeholders para facilitar el proceso evaluativo.* *Presenta una valoración notable en términos de credibilidad y exactitud en el diseño y los métodos aplicados (8,75), pues su diseño se corresponde con los intereses de la evaluación y se aplicaron diferentes métodos que cubren aspectos cuantitativos y cualitativos. En cuanto a la consideración de lo contrafactual, esta evaluación no lo contempla, si bien se considera que no aplica teniendo en cuenta los intereses evaluativos.* *En cuanto a los procedimientos y criterios de selección aplicados (5) no se explicitan en todos los casos y en el caso de la encuesta presenta insuficiente información para juzgar la bondad de la muestra.* *En cuanto al análisis de datos obtiene una calificación intermedia (5), pues triangula la información que se recaba a través de diferentes técnicas y lo expone con claridad y hace un uso de la información de monitoreo. Le penaliza en este sentido que sólo a veces cuantifica las respuestas que se exponen (teniendo que recurrir a los anexos en el caso de la encuesta) y no recurre a fuentes externas para la comparación durante el análisis (sólo algo para contextualizar). Y, finalmente, presenta credibilidad y exactitud en el tratamiento y enjuiciamiento de los criterios de evaluación (9) que le son de aplicación (a excepción de los relativos a impactos y eficiencia).*
	18,8	*Calidad del informe y utilidad:* [Sobre 10]	*La evaluación de la calidad y utilidad presenta resultados notables (9,38), pues el informe da respuesta adecuada a todas las preguntas de evaluación, describe de forma balanceada sus fortalezas y debilidades, diferencia moderadamente valoraciones descriptivas y valorativas y existe trazabilidad entre hallazgos, conclusiones y recomendaciones. Las recomendaciones que emite son relevantes y su grado de concreción adecuado (reforzado por sus detallados plan de acción y estrategia de socialización), junto a su articulación en una propuesta de diseño para el nuevo PND. Además contempla un resumen ejecutivo como parte del informe.*
	5,3	*Exhaustividad y adecuación de la documentación:* [Sobre 10]	*En el criterio de exhaustividad y adecuación de la documentación la evaluación recibe una calificación suficiente (6,5), en tanto que no presenta en anexos los TdR y en estos sólo incluyen algunos de los documentos y datos que especifican el trabajo de campo realizado, aunque si anexa los instrumentos que aplica. Presenta una estructura adecuada y su lectura está facilitada por gráficos y tablas aunque no con resúmenes de los capítulos*

0.9 Resumen de fortalezas de la evaluación.
• *Detallado Plan de acción y concreta y específica Estrategia de socialización como instrumentos facilitadores de la aplicación de las recomendaciones de esta evaluación.* • *Profundo y detallado examen técnico del diseño metodológico del PND en términos de racionalidad, pertinencia y coherencia interna y externa.* • *Importante trabajo de campo cualitativo con una buena sistematización de información que se expone con claridad para los lectores lo cual debe maximizar su utilidad final, facilitando su aplicación.* • *Exhaustivo trabajo de triangulación de la información, asentando en la introducción preguntas comunes a través de diversas técnicas, expuesto con claridad y de fácil interpretación.*
10 Resumen de debilidades de la evaluación.
• *Concede protagonismo en la triangulación al criterio experto del equipo evaluador (una de las cuatro fuentes trianguladas) sin exponer con claridad los hallazgos que lo justifican ni las prácticas internacionales que se refieren como base para su juicio (salvo excepciones). Si puede constituir una fuente relevante el dictamen experto que realizan sobre la calidad de la planificación si bien esto no se presenta así ni tampoco está sistematizado de manera que el lector pueda encontrar las fuentes y la trazabilidad de sus valoraciones. Esto resta rigor técnico al juicio final de la evaluación.*

- *Prima en exceso la claridad y facilidad de uso y lectura del informe, descuidando aspecto técnicos y metodológicos relevantes, como ir presentando los datos que sustentan los hallazgos de la evaluación, lo cual exige al lector a un tratamiento y revisión de los anexos para conocer las bases que sustentan el juicio de valor que se emite (esto ocurre especialmente en la encuesta), descuidando también el hecho de ir referenciando continuamente la fuente de la que se nutre los datos y hechos que se exponen.*

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	Se trata de una evaluación externa.
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	5	
		Evaluación externa (100% externo a instituciones evaluadas)	**10**	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y análisis de datos y presentación de resultados	Ninguna / No aplica	0	Se eligieron fuentes primarias y secundarias. Como fuentes primarias se contemplan los *stakeholders* a través de diferentes técnicas cuantitativas y cualitativas; y como fuentes secundarias la información relativa al PND y el marco normativo que puso a disposición MIDEPLAN, junto a experiencias internacionales. Estas fuentes se aplicaron en el levantamiento de datos y se usan en su análisis.
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	5	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas y se aplican en todas las etapas	**10**	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos? **Fases**: planificación e implementación.	Ninguna / No aplica / No hay información en los documentos consultados	0	*No hay referencias expresas a la implicación activa de los grupos meta y otros grupos interesados en la planificación e implementación de la evaluación.* *En la implementación, las referencias más directas que se realizan son: que hubo participación de los agentes implicados en la consulta realizada en la evaluación y que hubo una validación permanente (validación de los resultados preliminares con las autoridades de MIDEPLAN, Ministerio de Hacienda y de la Controlaría General de la República) y puntos de control sobre todos los documentos entregados con las partes interesadas.* ***Grupos implicados**: funcionarios públicos involucrados en el proceso y representantes de la sociedad civil y sector privado interesados en el PND 2011-2014). Entre el sector público destaca: MIDEPLAN, enlaces sectoriales de Ministerios Rectores del Sector, Presidencia, Ministerio de Hacienda, Contraloría, Comisión de Ingreso y Gasto Público de la Asamblea Legislativa. También participan en encuestas Ministerios, Órganos Adscritos a Ministerios, Autónoma o Semiautónoma, etc.* *A través de grupos focales participaron otros actores específicos: Estado de la Nación, CADEXCO, UCR, CICAP y BID; Defensoría de los Habitantes, Municipalidad de San José, AECID, FEMETROM, Consultor Independiente, Alcalde de Upala, Alcaldesa de Orotina, UNGL y MIDEPLAN; GIZ/Programa NEL, MINAE, SETENA, IMN, ICE, Consultor local GIZ, AyA y SENARA.*
		No hay participación	1	
		Participan escasa o desigualmente en planificación o implementación	**5**	
		Participan activamente en la planificación y en la implementación	10	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	*Las preguntas de evaluación fueron formuladas de forma clara. También fueron adaptadas al contexto pues hay concepciones particulares ad hoc de los diferentes criterios evaluativos, como en el caso de la eficacia. Es particular la formulación de la última pregunta evaluativa: Pues en realidad lo que pide es que las recomendaciones de la evaluación del PND 2011-2014 se orienten a definir una propuesta de mejora para el PND 2015-2018. Dada la particularidad de esta evaluación, es pertinente esta apreciación para evitar equívocos en el encargo.* ***Las preguntas de evaluación (y los criterios a los que se asocian) son:*** ***Pertinencia**: ¿En qué medida cumple el Plan Nacional de Desarrollo 2011-2014 con su objetivo de orientar y priorizar el quehacer institucional hacia metas nacionales de desarrollo? ¿Por qué?*
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	5	
		Las preguntas son claras y adaptadas	**10**	

				***Eficacia**: ¿Cuál es el grado de evaluabilidad del PND 2011-2014 según su estructura de formulación (considerando sus objetivos, indicadores y metas en todos los niveles)?* ***Calidad de la implementación**: ¿Qué elementos legales, presupuestarios y operativos se necesitan para garantizar que la ejecución del PND tenga carácter vinculante y obligatorio para las instituciones públicas?; ¿Cuáles son los procesos que representan debilidades, obstáculos o amenazas en la elaboración del PND 2011-2014?, ¿Por qué?* ***Sostenibilidad**: ¿Cómo intervienen los elementos de negociación y diálogo político entre jerarcas sectoriales, funcionarios técnicos y otros actores durante la construcción de propuestas sectoriales?* ***Transversal a los 4 criterios**: ¿Cuáles son las fortalezas más importantes detectadas en el proceso de elaboración del PND 2011-2014?, ¿Cómo favorecen estas fortalezas a la ejecución del PND?* ***Propuesta**: ¿Cuáles son las opciones e instrumentos de mejora que pueden incidir en la elaboración de futuros PND?*
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica**	**0**	*No se dispone de los TdR entre la documentación de base para esta metaevaluación.*
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	Según se desprende de la lectura del informe no se refiere la realización de un taller de inicio con stakeholders para facilitar su participación. Por el contrario se refiere la realización de tres misiones cortas del equipo de evaluación a Costa Rica completadas por trabajo a distancia en las que se despliegan entrevistas y reuniones con stakeholders.
		No hubo un taller de inicio con stakeholders para facilitar su participación	**1**	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	
		Hubo un taller de inicio con stakeholders	10	
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica	0	Enlace web al Informe de evaluación. (Presenta errores el enlace que no llegan a descargar el documento) https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/f7d32aa2-0948-412c-b0f6-2b6148ae5ed7/Evaluacion%20dise%C3%B1o%20metodologico%20PND%202013.pdf
		No son accesibles para el público (ni Informe ni resumen)	1	
		Son accesibles para el público (Resumen o extracto)	5	
		Son accesibles para el público (Informe y resumen)	**10**	
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**			
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	**Los objetivos están descritos de manera precisa:** **Objetivo general**: Mejorar el diseño metodológico y el proceso de elaboración del Plan Nacional de Desarrollo (PND) 2015-2018 a partir de la experiencia del PND 2011-2014 y proponer una nueva metodología para orientar la planeación hacia resultados más estratégicos y evaluables, que permita una vinculación con marcos presupuestarios de corto-mediano plazo y facilite una participación más amplia de actores sociales diversos y representativos. Los **objetivos específicos** a alcanzar con este mandato son: 1. *Identificar los elementos del contexto* nacional, normativo e institucional que determinaron el proceso de diseño metodológico y elaboración del PND 2011-2014; 2. *Analizar la incidencia de los distintos actores participantes* durante los momentos de negociación política y técnica que determinaron el proceso de diseño metodológico y elaboración del PND 2011-2014 y su resultado final; 3. *Analizar los factores que determinan la vinculación entre el PND 2011-2014 y los marcos presupuestal*es de corto y mediano plazo existentes para el mismo período;
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	
		Se describe el objetivo de la evaluación de manera imprecisa	5	
		Se describe de manera precisa el objetivo	**10**	

				4. *Analizar la concordancia de los objetivos del PND con las necesidades de su formulación y evaluabilidad*, para orientar las prioridades de desarrollo en el país, durante el periodo 2011-2014; 5. *Proponer una metodología del proceso de elaboración y diseño del PND* para futuros planes.
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica No se describe el objeto de evaluación **Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes** Se describe con precisión (al menos los ítems contemplados)	0 1 **5** 10	En el capítulo 1 "Contexto, objetivos y estructura del informe" se describe someramente el contexto que justifica la evaluación, y en cierta medida se apuntan algunos aspectos sobre el objeto evaluativo, si bien no se dedica un apartado específico a delimitarlo con precisión y caracterizarlo (grupos interesados, dimensión regional, volumen financiero, duración y niveles de intervención,…).
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholders)?	Ninguna / No aplica **No se realizó un análisis de stakeholders** Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.) Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	0 **1** 5 10	
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	**Ninguna / No aplica** No se elaboró Se elaboró pero no se usa en la evaluación o no es concluyente Se elaboró, se usa en la evaluación y es concluyente	**0** 1 5 10	No aplica a esta evaluación. Puede decirse que lo que se construye en la propuesta que presenta esta evaluación es una Teoría de la planificación (Basada en el Enfoque de Gestión para Resultados) . En cierto modo podría decirse que puede no aplicar, aunque lo cierto es que se puede construir a partir de la estructura de objetivos que presenta el PND.
	¿Se formularon hipótesis de impacto concluyentes?	**Ninguna / No aplica** No se formularon Se formularon pero no son concluyentes Se formularon y son concluyentes	**0** 1 5 10	No aplica a esta evaluación. Los motivos son los mismos que los descritos en el ítem anterior.
	¿Se diferencia entre los diferentes niveles de impacto?	**Ninguna / No aplica** No se diferencia entre niveles de impacto Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto) Se diferencia con claridad los diferentes niveles de impacto	**0** 1 5 10	No aplica a esta evaluación.
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica No se describe el contexto de la evaluación **Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención** Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	0 1 **5** 10	En el primer capítulo se presenta el contexto que justifica la evaluación ("1. Contexto, objetivos y estructura del informe"), si bien no se puede decir que haya un vínculo con la teoría de la intervención, algo que en esta evaluación no aplica directamente.
3.2	**Diseño y métodos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (Los TdR no explicitan los intereses de la evaluación) El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe) El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	0 1 5	No se dispone de los TdR de forma que permita formarse una opinión objetiva sobre los intereses de la misma, lo cual supone una limitación importante para emitir este juicio. Según se desprende de los objetivos y preguntas evaluativas que contiene el informe, el diseño de la evaluación se corresponde con los intereses que estos encierran: utilizar la evaluación del PND 2011-2014 para mejorar el diseño, implementación y resultados del PND 2015-2018. Articular la evaluación a partir del análisis documental y la consulta a los actores implicados en sus diferentes niveles orgánicos, territoriales y sectoriales, haciéndolo además a través de la triangulación de la información recabada mediante diferentes técnicas cuantitativas y cualitativas, se considera que se corresponde con sus

		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	**10**	intereses, siendo pues un diseño adecuado. Sobre todo, estructurar las recomendaciones en torno a una propuesta de metodología para el diseño es especialmente adecuado.
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica No se explicitan ni justifican el diseño y los métodos aplicados **Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados** Se explicitan y justifican suficientemente el diseño y los métodos aplicados	0 1 **5** 10	El diseño de evaluación se encuentra suficiente explicado y justificado. Se explican y justifican los métodos empleados. No obstante, su exposición se encuentra diluida a lo largo del informe, no presentando un apartado o subapartado específico dentro de la metodología destinado a justificar sistemáticamente cada una de ellas. Es por ello que se le otorga una calificación intermedia en este ítem.
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	**Ninguna / No aplica** No se consideró lo contrafactual Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró Se consideró lo contrafactual y se especificó cómo	**0** 1 5 10	No se contempló lo contrafactual. Puede considerarse que no aplica a esta evaluación.
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica Los métodos aplicados no cubren aspectos cuanti ni cuali Los métodos aplicados cubren aspectos cuanti o cuali **Los métodos aplicados cubren aspectos cuanti y cuali**	0 1 5 **10**	Se aplicaron diferentes métodos de levantamiento de datos que cubren aspectos cuantitativos y cualitativos.
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) Entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica Se aplicó sólo un método Se aplicaron dos métodos **Se aplicaron una pluralidad de métodos (más de dos)**	0 1 5 **10**	***MÉTODOS MIXTOS:*** ***Técnicas de investigación*** - ***Análisis documental (****Revisión documental de material bibliográfico entregado por MIDEPLAN****)*** - ***Entrevistas a profundidad*** *con actores clave: 33 participantes (Entre otros MIDEPLAN, enlaces sectoriales de Ministerios Rectores del Sector, Presidencia, Ministerio de Hacienda, Contraloría, Comisión de Ingreso y Gasto Público de Asamblea Legislativa y representantes de Sociedad Civil -ver el Anexo 8-.* - ***Taller de discusión de resultados preliminares con funcionarios:*** *36 participantes* - ***Encuesta on line*** *a funcionarios públicos: 196 universos / 128 participantes (65,3% tasa respuesta)* - ***Grupos focales*** *sobre temas estratégicos para la metodología de elaboración del PND 2015-2018: 31 participantes en total:* 1. *Participación y validación de la sociedad civil; 7 participantes (Sociedad civil)* 2. *Desagregación regional de las políticas; 14 participantes (Regiones y municipios)* 3. *Transversalidad del tema ambiental en varios Ejes de Acción; 10 participantes* ***Análisis de datos:*** - ***Herramientas:*** *Enfoque de Gestión para Resultados.* - ***Estrategia:*** *triangulación y retroalimentación con actores clave* - ***Otros análisis realizados.***
3.3	**Procedimiento de selección**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección **Se explicitan en algunos casos (técnicas)** Se explicitan los procedimientos y criterios de selección	0 1 **5** 10	En la Pág. 6 expone una justificación de la encuesta y grupos focales, pero no puede decirse que haya una justificación expresa y completa de los procedimientos y criterios de selección de todas las técnicas aplicadas a lo largo del informe ni en el capítulo específico de metodología. En la encuesta, que es la que más desarrolla este aspecto, se omiten datos relevantes para poder juzgar su validez como sustento de las conclusiones y recomendaciones.
3.3.2		Ninguna / No aplica	0	***Encuesta on line: [****Muestra* de 196 funcionarios (Responden 128; 65,3% de respuestas]

	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? Se entiende que aplica esta pregunta para las Encuestas	La muestra no es adecuada **La muestra no es adecuada o el error no es aceptable** La muestra sí es adecuada y el error aceptable	1 **5** 10	*En el informe se indica que* "Se definió que … la muestra debía estar por encima de 100 para garantizar un número de encuestas que permita garantizar los niveles de confiabilidad del resultado. Y, como estrategia para evitar el riesgo de no retorno, que por lo general se presentan en las encuestas on-line, se consideró una tasa de no retorno de 40% y se incluyó un número adicional de entidades y funcionarios en la muestra (es decir 78 encuestados adicionales)". *No se indica expresamente en los apartados dedicados a exponer la metodología bajo qué criterios o procedimientos se obtiene este número prefijado como mínimo aceptable para la muestra (100), y dado que tampoco indica expresamente en estos apartados el marco muestral o la población a la que se dirige la encuesta, no es posible conocer su representatividad o "qué tan extensas fueron las muestras" pues no se conoce si estamos ante un muestreo probabilístico o no probabilístico; intuyéndose que se trata de* un muestreo no probabilístico a una muestra de expertos pues no todas las unidades tienen tampoco la misma probabilidad de ser elegidas. Incluso no se conoce con claridad si se administra el cuestionario a toda la población que aplica. No obstante, si se atiende a lo que se apunta en otros apartados del informe, cabría pensar que el cuestionario aplicado se dirige a toda la población de entidades y órganos implicados en el PND 2011-2014. Esto es así pues en la página 21 se indica que el Estado de Costa Rica está conformado por 329 entidades y en la página 18 que "... fueron identificados por MIDEPLAN 100 entes y órganos como universo de entidades a las cuales aplican las líneas estratégicas del PND 2011-2014". En cuanto a los procedimientos de selección aplicados, en el informe se indica que "se previó que máximo hasta el 26% de la muestra podría ser MIDEPLAN para no sesgar el resultado. Ese número se definió también pensando en poder consultar ampliamente a las regionales de MIDEPLAN, y teniendo en cuenta que a diferencia de las características de otras entidades, muchas más unidades dentro de la entidad participaron en la elaboración del PND". Finalmente se indica que "A partir de los cargos directivos claves, enlaces sectoriales y enlaces institucionales seleccionados, se les invitó a presentar las personas que cumplieran las características definidas en sus entidades": **Cuadro 1: Estadísticas de recopilación de datos de la encuesta** — **Total/Razón** Invitaciones enviadas — 196 / 196 Declinaron participar — 3/196 Encuestas completadas — 128/196 Respuestas incompletas — 31/196 Por los motivos expuestos se le otorga una calificación intermedia en este ítem, no tanto por lo inadecuado de la muestra sino por la falta de información relevante al respecto para poder juzgarlo.
3.4	**Análisis de datos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali Se triangulan datos cuanti-cuali pero no de manera generalizada **Se triangulan datos cuanti-cuali de manera generalizada**	0 1 5 **10**	Si se triangulan datos cuanti-cuali. La exposición y uso de la triangulación en el análisis se ve facilitada por la presentación de tablas en las que se indica la valoración final obtenida de cada cuestión que se evalúa a partir de cada técnica., por lo que puede decirse que si aplica una triangulación en sentido estricto y no una generalización de combinación de técnicas para la contrastación de información.
3.4.2	¿Se cuantifican las respuestas siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	Ninguna / No aplica **No se cuantifican las respuestas cuando es posible** Se cuantifican las respuestas a veces cuando es posible Se cuantifican las respuestas de manera generalizada cuando es posible	0 **1** 5 10	No es habitual este recurso en el informe y su uso puede considerarse anecdótico. Frente a esta opción –cuantificar respuestas cuando es posible- para facilitar la lectura contextualizada de la evaluación se optó por presentar para cada pregunta y variable analizada el resultado global obtenido en cada técnica aplicada (encuesta, taller y grupos focales). No obstante, se presenta en anexo el detalle de las respuestas de la encuesta y los principales resultados de las restantes técnicas cualitativas. Esta opción del anexo exige un tratamiento adicional al lector para conocer los porcentajes que aplicarían a cada caso.
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica No se usaron datos del monitoreo	0 1	Los principales documentos y datos de esta naturaleza que se consultaron se presentan en el Anexo 1 del Informe bajo la categoría de "Documento/Estudios/Planes"

		Se usaron datos del monitoreo de manera insuficiente	5	No obstante, al tratarse de una evaluación de diseño el uso en sí de los datos de estos informes no es un recurso frecuente en el informe, no tanto por insuficiente, sino porque su uso no es que pueda considerarse que no aplica sino que no se erige en una fuente clave, como sí lo sería para evaluar la implementación o la eficacia (que aquí se concibe en términos de evaluabilidad). Así, son contadas las referencias explícitas a este tipo de datos, como las que se efectúan sobre: ▪ Informe Anual de Cumplimiento de Metas PND 2012 (MIDEPLAN). Citado en las páginas 30 y 32 ▪ Análisis sobre el cumplimiento de metas del PND 2011-2014 (CGR) y Análisis sobre el cumplimiento de metas del PND 2011-2014 (CGR, 2012). Citados, por ejemplo, en las páginas 20 y 29
		Se usaron datos del monitoreo de manera suficiente	**10**	
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica	0	No es habitual este recurso en el informe y su uso puede considerarse anecdótico. Puede justificarse por el tipo de evaluación de diseño el escaso uso de este tipo de datos o fuentes externas, aunque es cierto que es habitual que los análisis de pertinencia los utilicen para delimitar o ilustrar el tipo de necesidades de la población, territorios o sectores. En el informe se habla por ejemplo de las diferentes necesidades de regiones en determinados indicadores pero no se acompaña de datos de fuentes externas.
		No se recurrió a fuentes externas	**1**	
		Se recurrió a fuentes externas de manera insuficiente	5	
		Se recurrió a fuentes externas de manera suficiente	10	
3.5	**Criterios de evaluación**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa	El informe se estructura en su tercer apartado a partir de la respuesta a las preguntas evaluativas, articulando para cada una de ellas su respuesta a partir de los resultados globales de las técnicas aplicadas, entremezclando datos descriptivos con interpretaciones y juicios valorativos por parte del equipo evaluador. Llama la atención que se aporta como una de las técnicas el juicio del equipo evaluador "a partir de las buenas prácticas internacionales" según se indica en el informe, si bien no se presenta en ningún momento cuáles son dichas prácticas en las que se apoya, quedando como una suerte de dictamen experto del equipo evaluador no basado en hallazgos del proceso evaluativo[39]. Esto no constituye en rigor una fuente de información para la evaluación pues no permite seguir la trazabilidad del enjuiciamiento a partir de los datos cuantitativos y cualitativos recabados. Tiene además una importante incidencia en el juicio final ya que es una de las cuatro fuentes que triangula para emitir el juicio final.		
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	La pertinencia constituye uno de los criterios de evaluación contemplados, dedicando concretamente el punto 3.1 a la exposición de los hallazgos al respecto "3.1 Análisis sobre la Pertinencia del PND 2011-2014"
		No lo expone -1	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	5	
		Lo expone de forma concluyente y comprensible	**10**	
	b) logró las metas definidas	Ninguna / No aplica	0	La eficacia constituye uno de los criterios de evaluación contemplados, dedicando concretamente el punto 3.2 a la exposición de los hallazgos al respecto "3.2 Análisis sobre la Eficacia del PND 2011-2014". Cabe apuntar que más que concebir la eficacia como nivel de logro de los objetivos previstos, se concibe como el grado de cumplimiento del objetivo del PND en lograr ser instrumento de seguimiento y evaluación, y fuente de información para la retroalimentación continua de las acciones del Gobierno. En este sentido asocia esta respuesta a la pregunta sobre la evaluabilidad del PND
		No lo expone -1	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	c) es eficiente	**Ninguna / No aplica**	**0**	No constituye un criterio de evaluación la eficiencia.
		No lo expone	1	

[39] Son contados los casos en los que se puede inferir una imprecisa referencia a casos o experiencias internacionales. En general estos no se exponen como hallazgos de la evaluación sino que sus contadas referencias (concretamente 4) se entremezclan en la propia exposición de propuestas de mejora (recomendaciones).

Entre ellas se encuentran la que se efectúa en la página 71 "…En este sentido, algunos países han optado, para mejorar la articulación plan-presupuesto y darle más operatividad al plan, la solución institucional "de un sombrero" es decir unificar en un ministerio las funciones de planificación y presupuestación, tal es el caso de Brasil o Corea del Sur, solo para mencionar dos casos ..." También en la página 82: "Colombia y Brasil son dos de los países en que esta práctica ha consolidado la planificación. Más recientemente República Dominicana aprobó su estrategia de mediano plazo por ley del Congreso...:". En la página 85: La primera es buscar fuentes de financiación de las evaluaciones desde la programación misma de los programas y proyectos como sucede actualmente en Chile, México y Colombia… Incluso, en algunos países se discuten las evaluaciones en las sesiones de aprobación del presupuesto con los Congresos.". O, la nota al pie de la página 58 sobre el Fortalecimiento de las capacidades para la gobernabilidad PNUD. Otras referencias son las que se aportan en el plan de socialización (página 110 sobre estándares internacionales).

		Lo expone pero no de forma concluyente y comprensible	5	Simplemente en la propuesta de diseño que contiene la evaluación entre sus recomendaciones se propone este tipo de foco evaluativo: "Eficiencia en la implementación del PND":
		Lo expone de forma concluyente y comprensible	10	
	d) causó impactos (outcome e impacto) deseados y no deseados.	**Ninguna / No aplica**	**0**	No constituye un criterio de evaluación el impacto por lo que no puede decirse que el informe los exponga con claridad o sean concluyentes. No obstante, puede decirse que el informe si expone con claridad efectos deseados y no deseados que se derivan de la aplicación del PND 2011-2014 pero en el marco de otros análisis evaluativos y no el propio y estricto del impacto.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	e) causó impactos duraderos	Ninguna / No aplica	0	Podría decirse que los impactos duraderos están asociados a la sostenibilidad, que constituye uno de los criterios de evaluación contemplados, dedicando concretamente el punto 3.4 a la exposición de los hallazgos al respecto "3.2 Análisis sobre la sostenibilidad del PND 2011-2014". No obstante se trata de una concepción particular de la sostenibilidad, muy adaptada al contexto específico de esta evaluación del diseño metodológico del PND 2011-2014. Para evaluar el criterio de sostenibilidad, se analizan si existen los mecanismos institucionales para dar credibilidad a los procesos asociados a la validación de los contenidos del PND y garantizar el nivel de compromiso institucional suficiente para hacer creíble las metas fijadas para el PND y la rendición de cuentas permanente al interior y al exterior del gobierno como mecanismo de responsabilización (accountability). Por ello, la sostenibilidad del PND se mide en dos ámbitos: i) los mecanismos legítimos de validación política y técnica y ii) los mecanismos para garantizar la responsabilización por parte de los funcionarios en el contexto del M&E.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible / o aunque se aborda y puede considerarse concluyente no constituye un foco central de la evaluación	**5**	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	El informe de evaluación analiza la información disponible sobre avances del PND a partir de los diferentes informes publicados al respecto y analiza los diferentes sistemas de información de las instituciones implicadas en su análisis sobre la evaluabilidad del PND 2011-2014. Presenta conclusiones al respecto y aspectos de mejora para el PND.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	El informe de evaluación identifica a través de las técnicas cualitativas las percepciones existentes acerca de cambios observados en diferentes dimensiones y trata de vincularlos al diseño y la metodología aplicada en el PND. Por tanto, puede decirse que los reconduce al proyecto, aunque con un sentido particular respecto al espíritu de esta pregunta, pues no lo aborda en términos concluyentes de atribución/contribución.
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	**5**	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	**10**	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica	0	Se dedica el tercer capítulo a presentar la respuesta a las preguntas de evaluación (no se dispone de TdR para chequear que se corresponden con todas las que se listan en estos), dedicando un apartado a cada pregunta. Se considera que el informe ofrece una respuesta adecuada a todas las preguntas de evaluación.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	El informe presenta de forma balanceada fortalezas y debilidades del PND 2011-2014, algo que se aprecia con claridad en el propio resumen ejecutivo, pues en la respuesta sintética a cada pregunta se exponen las muchas debilidades que se observan y dedica un apartado específico en respuesta a la pregunta específica de evaluación centrada en identificar sus fortalezas y cómo estas contribuyen a la ejecución del PND.
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3		Ninguna / No aplica	0	

	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	No se distinguen con claridad declaraciones descriptivas y valorativas **Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe** Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	1 **5** 10	El informe no presenta de manera diferenciada, como suele ser habitual y recomendable, la exposición de los hallazgos del proceso evaluativo, sus conclusiones y sus recomendaciones. Este proceder de exposición diferenciada es el facilita la diferenciación entre declaraciones descriptivas (que concentran los hallazgos), y declaraciones valorativas (que concentran las conclusiones y el juicio evaluativo que encierran). No obstante, se ha optado por ofrecer a cada pregunta lo que se denomina "respuesta sintética" en la que se entremezclan tanto hallazgos como conclusiones y recomendaciones, y por tanto, las declaraciones descriptivas y valorativas. Dicho esto, normalmente se suele identificar cuándo se está ante un tipo de declaraciones u otro. Por tanto, no puede decirse que no se distingan en el conjunto del informe estas valoraciones, sino que presenta elementos de mejora para facilitar al lector su clara distinción en el informe para el lector.
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica No se especifica Se especifica en ocasiones **Se especifica en general**	0 1 5 **10**	Todas las preguntas presentan en una tabla sintética el resultado final de la valoración que desde cada técnica aplicada se obtuvo sobre cada elemento o variable que se evalúa. Junto a ello, en la explicación y justificación suele citarse en qué técnicas, datos o análisis realizados se basan las valoraciones y estimaciones que se presentan.
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica Las conclusiones no están basadas en los resultados Alguna o algunas conclusiones no están basadas en resultados **Todas las conclusiones están basadas en los resultados**	0 1 5 **10**	El hecho de haya una respuesta a cada pregunta en la que se exponen tanto hallazgos, como conclusiones y recomendaciones, facilita examinar que efectivamente las conclusiones están basadas en los resultados.
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica no se llega ni de forma lógica ni comprensible se llega de forma lógica o comprensible pero no las dos **Se llega de forma lógica y comprensible**	0 1 5 **10**	El hecho de haya una respuesta a cada pregunta en la que se exponen tanto hallazgos, como conclusiones y recomendaciones, facilita examinar que efectivamente se llega de forma lógica y comprensible de las recomendaciones a las conclusiones.
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica No son relevantes ni específicas Son relevantes o específicas pero no las dos **Son relevantes y específicas**	0 1 5 **10**	Las recomendaciones son relevantes y específicas para los destinarios y concretas. Cabe destacar a este respecto que además de las propias recomendaciones que se van presentando para cada pregunta evaluativa (Capítulo 3), se presenta en el Capítulo 4 una propuesta integrada de mejora del PND 2015-2018 que articula todas aquellas recomendaciones bajo la GpR, presentando hasta 31 propuestas específicas y concretas en las que en muchas de ellas se identifican destinatarios concretos, algo que se refuerza con el plan de acción y la estrategia de socialización, también detallada por agente y recomendación/propuesta formulada.
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica El Informe no contiene un resumen El informe contiene un resumen estructurado **El informe contiene un resumen estructurado, bien legible/comprensible**	0 1 5 **10**	El informe comienza con un sintético resumen de 8 páginas, estructurado en torno a las preguntas de evaluación y expone también la propuesta de diseño del PND 2015-2018. Es legible y comprensible.
5.	**Exhaustividad y adecuación de la documentación**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica **Los TdR no son accesibles en anexos** Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción Los TdR son accesibles en la lengua común de trabajo o en una traducción	0 **1** 5 10	Los TdR no forman parte del cuerpo de anexos del Informe
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de	Ninguna / No aplica No se incluyen en anexos	0 1	En los anexos se incluye la relación de personas participantes y entidades representadas en las entrevistas, grupos focales y talleres de discusión de resultados preliminares. Y en la exposición de la metodología se indican

	lugares y organizaciones visitadas y de personas entrevistadas?	**Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos**	**5**	las fechas y lugar de le celebración del Taller con funcionarios. También se expone en metodología las fechas clave del lanzamiento de la encuesta online. No se incluye en anexos el cronograma sintético de la fase de campo con los lugares en los que se celebraron y sus fechas.
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	En Anexos se incluye: - Relación de la documentación analizada que fue entregada por MIDEPLAN, aunque no incorpora un apartado de bibliografía. - Cuestionario aplicado en encuesta online. - Guías de los grupos focales. Además de presentar fundamentalmente una síntesis de estos grupos, se expone de manera general las guías y fundamentos del diseño de estos grupos. No se incluyen en anexos las guías de entrevistas aplicadas ni la literatura utilizada, al margen de la documentación aportada por MIDEPLAN.
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	**10**	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	Ninguna / No aplica	0	El informe presenta una estructura adecuada. No obstante en la respuesta a las preguntas evaluativas (capítulo 3) se echa en falta una diferenciación entre hallazgos, conclusiones y recomendaciones. El hecho de que una pregunta evaluativa solicitase expresamente una propuesta identificar las mejoras para futuros PND concretada en una propuesta metodológica para el PND 2015-2018 hace que se visualicen en mayor medida éstas propuestas como recomendaciones de esta evaluación (propuestas como se las llama en este Informe), quedando en un segundo plano aquellas que se fueron formulando en el estudio evaluativo sobre cada pregunta, aunque muchas de ellas se recuperan en esta propuesta final. En cuanto a elementos que faciliten la lectura se contemplan normalmente tablas, en algunas ocasiones gráficos aclaratorios –no sobre representación de datos- y no se incorporan resúmenes de los capítulos.
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	**5**	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	10	

4.4.5 Evaluación 05: REDCUDI

0.1	Nombre del proyecto / programa evaluado	***EVALUACIÓN DEL DISEÑO, GESTIÓN Y EJECUCIÓN DEL BENEFICIO "ALTERNATIVAS DE ATENCIÓN Y DESARROLLO DE LA NIÑEZ", QUE ATIENDEN A LOS NIÑOS Y A LAS NIÑAS QUE RECIBEN EL SUBSIDIO DEL IMAS EN EL MARCO DEL PROGRAMA RED NACIONAL DE CUIDO Y DESARROLLO INFANTIL, PERÍODO 2014 Y 2015***
0.2	Institución que implementa (p.ej. Ministerio)	*Instituto Mixto de Ayuda Social (IMAS) y Secretaría Técnica de la Red Nacional de Cuido y Desarrollo Infantil*
0.3	Evaluadores	*Centro de Investigación y Capacitación en Administración Pública, Universidad de Costa Rica (CICAP-UCR)*
0.4	Monto del proyecto/ programa	*No se especifica. Las únicas referencias son: "En el 1er semestre 2015, el IMAS otorgó subsidios … para sufragar el costo de atención en Alternativas de Atención a la Niñez … para una inversión de ¢7.368.941,3 miles de colones"; y "la meta presupuestaria para 2015 ascendió a ₡21,721,039,200,00*
0.5	Monto para la evaluación	*¢50.000.000,00 (50 millones de colones) [Aproximadamente 85.600 €]*

0.6 Descripción del proyecto/programa evaluado (un párrafo)
El beneficio del IMAS consiste en la entrega de un subsidio a las familias para que sus hijos asistan a las Alternativas de Atención y Desarrollo de la Niñez (AADN), prioritariamente niños y niñas de cero a siete años, que viven en condiciones de pobreza, pobreza extrema y con vulnerabilidad social; no obstante, según necesidades específicas y disponibilidad presupuestaria, se pueden incluir niños y niñas hasta de doce años de edad que se encuentran en pobreza extrema y pobreza. Las transferencias económicas deben ir acompañadas por procesos o acciones de promoción y orientación individual o familiar, que promuevan la movilidad social ascendente. Y las modalidades de AADN ofrecen diferentes servicios y horarios de atención a padres, madres o encargados legales de niños y niñas prioritariamente en condiciones de pobreza son: Centros de Cuido y Desarrollo Infantil (CECUDI) - Oferta de servicio prestada por alternativas de cuido municipales-; Hogares Comunitarios - Oferta de servicios que presta una Madre Comunitaria-; y Centros infantiles - Oferta de servicio prestada por alternativas de cuido privadas o mixtas, incluidas cooperativas, asociaciones de desarrollo, asociaciones solidaristas y pequeñas empresas-
0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL (máximo media página)
Evaluación con apoyo de FOCEVAL mediante la modalidad de Padrinazgo. Según apuntan los TdR (Pág.15 y 16): *"Se contará con el apoyo de FOCEVAL a lo largo de todo el proceso de la evaluación, en donde se aportará la colaboración de un experto internacional en la figura de padrino el cual estaría contribuyendo a la calidad del proceso de evaluación, reforzando las capacidades de los actores implicados en dicho proceso.* *La modalidad del "padrinazgo de una evaluación" se basa en la figura de un consultor internacional experto en evaluación (padrino) que acompaña el proceso de evaluación asesorando técnica y metodológicamente al Equipo de Gestión y a través de éste, al Equipo Evaluador.* *Está previsto que el asesoramiento del padrino tenga lugar desde la etapa de diseño de la evaluación, de manera que pueda acompañar en la toma de decisiones clave para el resto del proceso: delimitación del objeto de la evaluación, definición las preguntas claves, diseño metodológico, elaboración de términos de referencia para contratar al equipo evaluador, etc. El apoyo del padrino se extenderá hasta concluida la etapa de diseminación de los resultados y la elaboración de una estrategia para la aplicación de las recomendaciones de la evaluación ("management response")".* *Al margen de esta delimitación de la modalidad que hacen los TdR, en el informe no se aprecia ninguna referencia expresa a las bondades o limitaciones que ha supuesto esta experiencia, ni directa ni indirectamente. Tan sólo cabe señalar que se apunta que la evaluación con una estructura de gestión (Equipo gestor y Equipo tomador de decisiones).*

0.8 Cualificación cuantitativa (con comentarios) según criterios			
67,5 [Sobre 100]	**7,5**	*Imparcialidad e independencia:* [Sobre 10]	Arroja buena puntuación en Imparcialidad e independencia (7,5) pues se trata de una evaluación externa en la que además se considera que las fuentes seleccionadas permiten capturar una pluralidad de perspectivas en sus distintas etapas.
	6	*Equidad, participación y transparencia:* [Sobre 10]	La evaluación obtiene una calificación suficiente en este criterio (6), pues participan activamente en el diseño grupos interesados –mediante el taller de diseño de la teoría de la intervención- (aunque no se aprecia la participación activa en la implementación más que como informantes). En cuanto a las preguntas de evaluación si bien son claras se consideran complejas (cada una encierra más de un criterio evaluativo), y los TdR presentan una estructura adecuada y una extensión considerable que hubiese requerido mayor clarificación del objeto o un índice entre otros aspectos. Y tanto el informe como el resumen ejecutivo que aquél contiene están accesibles en internet.
	30	*Credibilidad y exactitud:*	La evaluación obtiene una valoración suficiente en este criterio (6,2), pues está descrita la meta/objetivo de la evaluación, describe con precisión el objeto evaluado y describe relativamente el contexto evaluativo, presenta un completo análisis de los stakeholders y una detallada teoría de la intervención.

		[Sobre 20]	Presenta una valoración notable en términos de credibilidad y exactitud en el diseño y los métodos aplicados (7,2), pues su diseño se corresponde con los intereses de la evaluación (con la salvedad del abordaje de la ejecución del Beneficio) y se aplicaron diferentes métodos que cubren aspectos cuantitativos y cualitativos. En cuanto a la consideración de lo contrafactual, esta evaluación no lo contempla, si bien se considera que no aplica teniendo en cuenta los intereses evaluativos. En cuanto a los procedimientos y criterios de selección aplicados (5) no se explicitan en todos los casos y en el caso de la encuesta presenta una muestra adecuada con elevados errores en algunos casos. En cuanto al análisis de datos obtiene una calificación intermedia (5,25), pues triangula la información que se recaba a través de diferentes técnicas (aunque no generalizadamente) y cuantifica las respuestas obtenidas cuando es posible, incluso a veces en las técnicas cualitativas. Le penaliza en este sentido el casi inexistente uso de datos del monitoreo y de fuentes externas para la comparación durante el análisis. Y, finalmente, presenta un bajo registro en credibilidad y exactitud en el tratamiento y enjuiciamiento de los criterios de evaluación (4,4), penalizándole en los que le son de aplicación la falta de respuesta suficiente en términos de eficiencia y sistema de monitoreo del programa.
	18,8	*Calidad del informe y utilidad:* [Sobre 10]	La evaluación de la calidad y utilidad según las preguntas planteadas presenta resultados notables (9,38), pues el informe describe de forma balanceada sus fortalezas y debilidades, diferencia valoraciones descriptivas y valorativas y existe trazabilidad entre hallazgos, conclusiones y recomendaciones. Las recomendaciones que emite son relevantes y, en general, presentan un grado de concreción adecuado, aunque hay excepciones. Además contempla un resumen ejecutivo como parte del informe. Le penaliza el hecho de que alguna de las preguntas formuladas en los TdR no se atienen en esta primera fase de la evaluación.
	5,3	*Exhaustividad y adecuación de la documentación:* [Sobre 10]	En el criterio de exhaustividad y adecuación de la documentación la evaluación recibe una calificación intermedia (5,2), en tanto que no presenta en anexos los TdR y en estos sólo incluyen algunos de los documentos y datos que especifican el trabajo de campo realizado, aunque si anexa alguno de los instrumentos que aplica. Presenta una estructura adecuada y su lectura está facilitada por gráficos y tablas aunque no con resúmenes de capítulos

0.9 Resumen de fortalezas de la evaluación.

- *Desarrolla un meritorio trabajo en la reconstrucción de la Teoría de la intervención siguiendo el Manual de MIDEPLAN.*
- *Formula un completo análisis de los stakeholders.*
- *Contempla un amplio trabajo de campo que abarca una diversidad de perfiles muy adecuada para el objeto que se evalúa.*
- *Ofrece un amplio balance de percepciones de actores clave que puede ser de utilidad para la toma de decisiones*

0.10 Resumen de debilidades de la evaluación.

- *Siendo que la evaluación se define como de ejecución (además de diseño y gestión), no contar datos de monitoreo supone una limitación muy importante. Sólo se trabajaron las bases de datos al efecto de construir la muestra, pero no se explotó para ofrecer información sistematizada sobre las realizaciones físicas y financieras del Beneficio evaluado.*
- *No se presta atención al periodo que se evalúa según los TdR (2014 y primer semestre de 2015), no ofreciendo un solo dato agregado del mismo, sino sólo algún dato aislado de subperiodos*
- *Los TdR presentaron una alta exigencia en cuanto a preguntas (aunque fueron mejoradas de su formulación inicial, según se aprecia en sus anexos), lo que unido a la alta exigencia de productos, genera un encargo de amplio alcance que quizás hubiese requerido una mayor focalización.*
- *Los Anexos presentan información de utilidad pero no tienen un índice que ayude a localizarla, lo cual puede limitar su uso.*
- *En ninguna de las técnicas se contempló si quiera a modo exploratorio la participación directa de la infancia en línea con el enfoque de derechos que promueve esta evaluación. Siendo complejo, alguna muestra en este sentido podría haber supuesto un aporte relevante e innovador, en línea con algunas experiencias en este tipo de programas.*

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	Se trata de una evaluación externa realizada por el CICAP-UCR y gestionada por MIDEPLAN e IMAS mediante la conformación de un Equipo Tomador de Decisiones y un Equipo Gestor de la evaluación en los que también está representada la Secretaría Técnica de la Red Nacional de Cuido. Los integrantes del equipo evaluador son 100% externos a IMAS y Secretaría Técnica como órganos responsables del Beneficio y de la Red Nacional de Cuido
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	5	
		Evaluación externa (100% externo a instituciones evaluadas)	**10**	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y análisis de datos y presentación de resultados	Ninguna / No aplica	0	Se eligieron tanto fuentes primarias como secundarias, si bien priman las fuentes primarias siendo escasas las fuentes secundarias ▪ Fuentes primarias: Entrevistas semiestructuradas y talleres todos ellos con stakeholders; cuestionario autoadministrado y encuestas a 143 funcionarios, 65 cocineras, 111 directoras(es), 368 representantes de familias y 111 instalaciones ▪ Fuentes secundarias: Fuentes bibliográficas y normativas y algunas fuentes de datos para la contextualización. Es en el levantamiento de datos donde más se echa en falta haber hecho un mayor uso de información sobre la ejecución física y financiera del Beneficio en el periodo a evaluar (2014-2015), máxime al tratarse de una evaluación de diseño y ***ejecución***.
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	**5**	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas y se aplican en todas las etapas	10	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos? **Fases**: planificación e implementación.	Ninguna / No aplica / No hay información en los documentos consultados	0	El informe no refleja la involucración activa de los grupos meta u otros grupos interesados. *Según se apunta en el informe "se trata de una evaluación participativa, se aprecia que la intención si fue desplegar un proceso de esta naturaleza, pero no presenta evidencias ni referencias que indiquen o muestren una participación genuina de los stakeholders y grupos meta en el diseño de la evaluación (delimitación, preguntas,…).* *Lo más llamativo en el diseño es la participación en el diseño de la Teoría de la intervención, que puede considerarse parte del diseño. Partió de un enfoque de evaluación participativa, en donde el punto de vista de los actores estratégicos vinculados al funcionamiento de las AADN y de la REDCUDI fue fundamental en la definición del "deber ser" de la intervención evaluada.* *En esta actividad, fueron implicados los siguientes **grupos**: Equipo Gestor y Equipo Tomador de Decisiones y funcionarios y funcionarias del IMAS que participaron en el "Taller de Validación de la Teoría de la Intervención de las AADN", donde revisaron las cadenas de resultados elaboradas por el Equipo Consultor.* *Y, en la implementación, lo que se aprecia en general es que más que una participación en sentido genuino, los stakeholders participan de las diversas consultas realizadas a través de diferentes técnicas, como informantes y, pudiendo calificarse así como una evaluación consultiva.*
		No hay participación	1	
		Participan escasa o desigualmente en la planificación o implementación / o Participan activamente en la planificación o en la implementación	**5**	
		Participan activamente en la planificación y en la implementación	10	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	*Los TdR exponen 5 preguntas principales de evaluación:* Su formulación está adaptada al contexto evaluativo, si bien son complejas, pues varias de ellas encierran en realidad más de una pregunta (lo que se aprecia observando la diversidad de criterios evaluativos sobre los que tiene que informar cada una de ellas –hasta tres en una misma pregunta-). De esta manera encierran un amplio encargo evaluativo (alcance). Pero sobre todo, destacan otros aspectos de las preguntas que dificultan la focalización del encargo: - Hay diferencias entre las preguntas que aparecen en los TdR (Pág.13) y las que presenta el Informe final (Pág.14). Aunque se presenta una justificación (Pág.93 del Informe), no queda claro y, en todo caso, es llamativo el cambio de las preguntas que motivan la evaluación. - Las preguntas que se añaden en el Anexo 4 de los TdR, desagregan las preguntas principales y aunque presentan diferencias sobre las que exponen los propios TdR (Pág.13), llama la atención que no se tienen en cuenta en el Informe de evaluación, salvo tangencial y difusamente. Dicho esto sobre las complejidades que presentan las preguntas de evaluación, cabe apuntar atendiendo al Anexo 5 de los TdR que presentan "Toda las preguntas identificadas en la etapa de programación de esta evaluación", que a
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	**5**	
		Las preguntas son claras y adaptadas	10	

				pesar de todo el proceso de formulación de preguntas fue objeto de un importante trabajo de priorización sobre las inicialmente previstas.
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	Ninguna / No aplica Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción) **Los TdR presentan una formulación clara (estructura) o comprensible (redacción)** Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	0 1 **5** 10	Los TdR conforman un documento extenso que presenta una estructura que trata de exponer con claridad el complejo entramado institucional y programático en el que se inserta el Beneficio que se evalúa y su redacción es comprensible. Dicho esto, no es fácil para el lector formarse una idea intuitiva del objeto que se evalúa, pues los propios términos o conceptos clave (como Beneficio, Alternativa o Modalidad) hubiesen precisado una mayor explicación y claridad. En cuanto a la estructura, puede considerarse adecuada en términos generales si bien su extensión hubiese precisado un índice de contenidos y de anexos que facilitase su uso y lectura.
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica No hubo un taller de inicio con stakeholders para facilitar su participación **No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)** Hubo un taller de inicio con stakeholders	0 1 **5** 10	No se aprecia en el informe que se celebrase un taller de inicio con stakeholders según se concibe en este criterio de metaevaluación en términos de promover la equidad, transparencia y participación. Por el contrario si se realizó esta actividad en el marco del diseño de la teoría de la intervención, si bien con una finalidad específica diferente.
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica No son accesibles para el público (ni Informe ni resumen) Son accesibles para el público (Resumen o extracto) **Son accesibles para el público (Informe y resumen)**	0 1 5 **10**	Se entiende por accesibilidad su disposición a través de internet. En este sentido, en Internet es posible acceder al informe completo de esta evaluación, donde se incluye el resumen ejecutivo pero no los Anexos. Enlace web: https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/adf401f5-b261-4433-a503-90020bba745c/Red_De_Cuido_Informe_final_evaluacion_I_fase.pdf?guest=true
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**			
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica No se describe el objetivo de la evaluación (Ni TdR ni Informe) Se describe el objetivo de la evaluación de manera imprecisa **Se describe de manera precisa el objetivo**	0 1 5 **10**	Los objetivos de la evaluación quedan indicados en la pág. 12 de los TdR y se reproducen el Informe de evaluación (Pág. 92), si bien en este no aparecen los dos últimos objetivos específicos de los TdR. Los objetivos, al igual que las preguntas, se formulan integrando los propios criterios de evaluación. *Objetivo General: Evaluar el diseño, gestión y ejecución del beneficio "Alternativas de Atención y Desarrollo de la Niñez", que atienden a los niños y a las niñas que reciben el subsidio del IMAS en el marco de la Red Nacional de Cuido y Desarrollo Infantil, para sustentar la toma de decisiones oportunas en el mejoramiento de su gestión y el cumplimiento de sus objetivos.* *Objetivos específicos:* - *Analizar la coherencia y pertinencia de las modalidades Alternativas de Atención y Desarrollo de la Niñez, que atienden a los niños y a las niñas que reciben el subsidio del IMAS con la teoría del Programa Red Nacional de Cuido y Desarrollo Infantil (REDCUDI).* - *Analizar la pertinencia, eficiencia y eficacia de las modalidades de Atención y Desarrollo de la Niñez que atienden a los niños y a las niñas con subsidio del IMAS (Infraestructura, equipamiento, capacitación, recurso humano, costo y atención), en el marco del Programa Red Nacional de Cuido y Desarrollo Infantil.* - *Determinar la pertinencia, eficacia, eficiencia y sostenibilidad de los procesos relacionados con el subsidio económico (solicitud, trámite, estudio, otorgamiento y control) que brinda el IMAS para que los niños y las niñas reciban los servicios de las Modalidades Alternativas de Atención y Desarrollo de la Niñez, en relación con los procedimientos institucionales IMAS.* - *Determinar si la articulación intra e interinstitucional entre los diferentes actores involucrados en las modalidades de Alternativas de Atención y Desarrollo de la Niñez que atienden a los niños y a las niñas con subsidio del IMAS es pertinente, eficiente y sostenible según lo establecido en el marco del Programa Red Nacional de Cuido y Desarrollo Infantil.*

				- *Determinar como el seguimiento y evaluación de las modalidades de Alternativas de Atención y Desarrollo de la Niñez, que atienden a los niños y a las niñas que reciben el subsidio del IMAS, son pertinentes, eficientes y sostenibles con respecto al logro de los objetivos institucionales del IMAS.*
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica No se describe el objeto de evaluación Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes **Se describe con precisión (al menos los ítems contemplados)**	0 1 5 **10**	*De todos los ítems que se contemplan en esta pregunta es el volumen financiero el que no se expone en el informe, salvo algún dato no global ni sistemático del periodo evaluado (2014, 2015) que no permite formarse un opinión de la envergadura presupuestaria del programa, ni siquiera en un ejercicio presupuestario (pues son datos en fechas específicas (datos del primer semestre de 2015 por ejemplo). No obstante, dado que se examina en detalle la intervención en el marco de la exposición de la teoría de la intervención, se le ha otorgado la máxima puntuación en esta pregunta. Como aspecto de mejora, se encuentra la necesidad para el lector de haber expuesto de manera sintética al inicio del informe las claves de la intervención y de los principales términos y conceptos en los que se asienta para facilitar su mejor conocimiento desde el inicio del informe.*
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica No se realizó un análisis de stakeholders Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.) **Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)**	0 1 5 **10**	*Se realizó un mapeo de actores (Aptdo.3.3) en el que se exponen los actores principales vinculados con la gestión de las AADN, en el marco de la REDCUDI, sus funciones y roles dentro del proyecto. Asimismo, se presenta un análisis del nivel de influencia sobre las AADN y su capacidad de acción de actores; y un análisis del nivel de operación en las AADN de los diversos actores donde se ubicación en diferentes niveles (instancias públicas de apoyo, instancias ejecutoras u otras instancias que participan en la atención de la población infantil).* *El uso que se hace de este análisis en el informe se aprecia en la conformación de instrumentos de recogida de información y en las técnicas aplicadas, así como en los análisis que se realizan, especialmente en los relativos a procesos y a gestión del programa y de las AADN.*
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica No se elaboró **Se elaboró pero no se usa en la evaluación o no es concluyente** Se elaboró, se usa en la evaluación y es concluyente	0 1 **5** 10	*Se elaboró una teoría de la intervención, especificada para los Hogares Comunitarios, Centros Infantiles Privados, Centros de Cuidado y Desarrollo Integral (CECUDI) y para la Red Nacional de Cuido y Desarrollo Infantil (REDCUDI). Esta teoría se apoya en su diagramación mediante el recurso de las cadenas de resultados, en línea con lo especificado en los TdR y en el propio Manual de MIDEPLAN. Puede afirmarse que desde un punto de vista teórico o conceptual puede ser concluyente, pero no se apoya en base empírica que la avale, no se formulan indicadores como solicitaban los TdR y no se usa a lo largo de la evaluación como eje vertebrador de la misma ni para tratar de contrastarla a partir de la información recabada. Más bien constituye un producto aislado de la evaluación, por otra parte, como la solicitaban los propios TdR. Es más suelen formularse usando fórmulas como "podría derivarse… o .. podría implicar …, permitiría atender …o se espera qué …."*
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica No se formularon **Se formularon pero no son concluyentes o no son explícitas.** Se formularon y son concluyentes	0 1 **5** 10	No se formularon de manera explícita hipótesis de impacto como tales, pero en la representación gráfica de las cadenas de resultados y en la explicación en texto que las acompaña se enuncia cambios y efectos que se espera que se deriven de los productos que entrega el programa y de su buen aprovechamiento por parte de los grupos meta, y por tanto puede decirse que de una u otra manera, están presentes, si bien de forma implícita. Esto se concreta especialmente en el subapartado que describe la Teoría del impacto. Ejemplos ilustrativos de esto que se apuntan son: - *"El hecho de que las niñas y los niños reciban cuidado en los Hogares Comunitarios, podría implicar una disminución de la exposición de las personas menores de edad a eventuales situaciones de riesgo, tales como la delincuencia y la drogadicción."* - *„El impacto en la calidad de vida en los niños y las niñas se centra en que éstos mejoren sus posibilidades de ser incorporados (as) activamente dentro de la sociedad costarricense"* - *" ... se espera que las economías locales se dinamizan mediante la gestión de microempresas privadas dedicadas al cuidado, atención y desarrollo integral de las personas menores de edad, utiliza capital humano y de otra índole, de procedencia local."* - *esta alternativa de atención pretende que los niños y las niñas mejoren sus posibilidades de inserción en el mercado laboral en su vida adulta y con ello su nivel de ingresos.*
	¿Se diferencia entre los diferentes niveles de impacto?	Ninguna / No aplica No se diferencia entre niveles de impacto **Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)**	0 1 **5**	Si se diferencia entre niveles de efecto en el marco de las cadenas de resultados que se formulan. Concretamente se contemplan dentro de lo que se denomina "Teoría de impactos" y aparecen "Efectos" e "impactos". No obstante dentro del nivel de impacto no se establecen diferentes niveles o graduaciones, por lo que se le otorga una puntuación intermedia en esta pregunta.

		Se diferencia con claridad los diferentes niveles de impacto	10	
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica	0	En el apartado 3 "Contexto y origen de la intervención evaluada" se expone el contexto global y se especifican algunos aspectos del contexto específico de las AADN que se contemplan en el Beneficio que se evalúa. De esta exposición el lector puede inferir vínculos con la Teoría de la intervención. Seguidamente se expone la Teoría de la intervención que detalla los insumos, actividades, productos, efectos e impactos así como los supuestos en que se basan, aunque no detalla específicamente en este apartado el contexto específico vinculado a dicha Teoría de intervención. Es por ello que se otorga una puntuación intermedia.
		No se describe el contexto de la evaluación	1	
		Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención	**5**	
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	10	
3.2	**Diseño y métodos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (Los TdR no explicitan los intereses de la evaluación)	0	*El diseño de evaluación se considera adecuado con los intereses de la misma. La mayor limitación que aprecia, teniendo en cuenta que se trata de una evaluación de diseño, gestión y ejecución, es la falta de consideración de fuentes de información externa y datos de seguimiento para caracterizar y ponderar la población objetivo y sus necesidades (para apoyar la evaluación de la pertinencia) y documentar evidencias sobre la ejecución del programa (evolución de participantes según perfiles y tipo de AADN, etc.). Esta falta de datos hace que la consideración del periodo evaluado según los TdR tenga reflejo limitado en el diseño evaluativo, circunscrito sobre todo a la encuesta (al seleccionar familias de esos periodos).* *Según se indica en el Informe es esta una evaluación del tipo formativa desarrollada por medio de una metodología participativa y con un enfoque mixto (cualitativo y cuantitativo) por medio de un equipo multidisciplinario de expertos y asesores del CICAP-UCR. Asimismo se indica que se desarrolló desde el enfoque de evaluación conocido como Evaluación centrada en la utilización, el cual pretende que la misma disponga de un uso orientado a la toma de decisiones a nivel institucional y el desarrollo de capacidades en actores claves (Patton 2000).* *Incluyó la elaboración del plan de trabajo, la reconstrucción de las Teoría de la Intervención de cada Alternativa y de la REDCUDI, así como el análisis de la pertinencia, eficiencia y eficacia de los servicios y procesos de las Alternativas. Para lograrlo, se desarrolló un proceso valorativo que abarcó:* ▪ *Una revisión bibliográfica y documental sobre el diseño del programa y de las alternativas.* ▪ *7 talleres con diversos públicos meta.* ▪ *Entrevistas semi estructuradas con informantes clave (6).* ▪ *Cuestionario auto administrado a 6 personas del Área Programática de Bienestar Familiar del IMAS.* ▪ *Encuesta a 143 funcionarios, 65 cocineras, 111 directoras(es), 368 representantes de familias y 111 instalaciones.*
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	5	
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	**10**	
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica	0	El diseño aparece justificado en cuanto al enfoque evaluativo adoptado "*Evaluación centrada en la utilización" (Patton 2000). Se justifica en el Informe esta elección por alinearse con dos normas de evaluación contempladas en el Manual para el Diseño y Ejecución de Evaluaciones (MIDEPLAN, 2014): utilidad y orientación al aprendizaje.* *Y en lo metodológico hay algunas referencias a niveles de análisis (estratégico, de gestión y operativo) pero no una reflexión sobre las opciones de abordaje de la evaluación, aunque se apuntan algunas decisiones metodológicas relevantes (sobre las preguntas, sobre la muestra, entre otras).* *En cuanto a la encuesta más que justificar su idoneidad se explica que es un requerimiento de los TdR.*
		No se explicitan ni justifican el diseño y los métodos aplicados	1	
		Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados	**5**	
		Se explicitan y justifican suficientemente el diseño y los métodos aplicados	10	
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	Ninguna / No aplica	0	
		No se consideró lo contrafactual	**1**	
		Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró	5	
		Se consideró lo contrafactual y se especificó cómo	10	
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica	0	Si se aplicó una pluralidad de métodos que cubren aspectos cuantitativos y cualitativos.
		Los métodos aplicados no cubren aspectos cuanti ni cuali	1	
		Los métodos aplicados cubren aspectos cuanti o cuali	5	

		Los métodos aplicados cubren aspectos cuanti y cuali	**10**	
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) Entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica Se aplicó sólo un método Se aplicaron dos métodos **Se aplicaron una pluralidad de métodos (más de dos)**	0 1 5 **10**	*Métodos y técnicas de recogida de datos (con indicación de las personas a las que se consultó y el número de ediciones de cada una de ellas)* ▪ **Revisión bibliográfica y** documental (más de 80 documentos) ▪ **Talleres con diversos públicos meta**: **7** (Públicos: destacan enlaces regionales de Infancia de ARDS, funcionarios/as vinculados a la gestión de la Secretaría Técnica REDCUDI durante 2011-2015, funcionarios/as de ARDS vinculados con la ejecución de las AADN y del Área Programática de Bienestar Familiar vinculados con la gestión del beneficio. ▪ **Entrevistas semi estructuradas: 13** (hay una discrepancia: en la página 25 se indican 15 entrevistas; en la página 101 se indican 13 y en la página 14 y en la figura 3 se indican 6 entrevistas).. ▪ **Cuestionario auto administrado**: **6** (Área Programática de Bienestar Familiar del IMAS). ▪ **Encuesta**: **143** funcionarios, **65** cocineras, **111** directoras(es), **368** representantes de familias y **111** instalaciones. Lo que da un total de **589** personas encuestadas y 111 instalaciones **Análisis** ▪ El proceso consideró la aplicación de diversas técnicas que arrojaron 3,000 páginas de resultados estadísticos, que propician cruces específicos de variables para cada criterio de evaluación. Esos resultados son producto de haber realizado 721 preguntas, en 5 instrumentos que fueron aplicados en total a 815 personas entre Funcionarios(as), Directores(as), Familias, Nutricionistas e Infraestructura, las cuales responden a 5 indagatorias diferentes. ▪ Se utilizó el contraste de información como estrategia de análisis. Entendida como recolección de información de diversas técnicas para responder a cada pregunta evaluativa. También se apunta el uso de la triangulación al contener algunas de las diferentes técnicas grupos de preguntas y cuestiones comunes.
3.3	**Procedimiento de selección**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección **Se explicitan en algunos casos (técnicas)** Se explicitan los procedimientos y criterios de selección	0 1 **5** 10	*Con respecto a la información cuantitativa relativa a las encuestas se explicitan en detalle los procedimientos y criterios de selección aplicados.* *Con respecto a la información cualitativa no se explicitan en detalle los procedimientos y criterios de selección aplicados.*
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? Se entiende que aplica esta pregunta para las Encuestas	Ninguna / No aplica La muestra no es adecuada **La muestra no es adecuada o el error es elevado** La muestra sí es adecuada y el error aceptable	0 1 **5** 10	Se realizaron: • 3 encuestas probabilísticas "generalizables a la población", Familias, Directores e Infraestructura • 2 encuestas de opinión aplicadas a Nutricionistas o cocineros y a los funcionarios(as). Las muestras son representativas en las encuestas probabilísticas pero en algunos casos los errores se consideran muy elevados. Los procedimientos de selección aplicados fueron los siguientes: Proceso de selección y análisis del componente Familia • Encuesta Probabilista • Muestra sistemática y estratificada por área regional y por tipo de alternativa generalizable a la población total • Tamaño de Muestra Propuesto: N = 463, Muestra Final N1= 368 • Tasa de Respuesta: 79.5 % • Porcentaje de No respuesta: 20,5% De estos una mayoría se debe a que no llego a la entrevista y/o ya no tenía al niño en el centro. • Sub Universo Expandible: 2617 Error de Muestreo: 4.6% • Universo Total: 8724 Error de Muestreo: 9.3% **Proceso de selección y análisis del componente Infraestructura** • Encuesta Probabilista • Tamaño de Muestra Propuesto: N = 121, Muestra Final N1= 111 • Tasa de Respuesta : 91.73% • Error de Muestreo: 8 puntos porcentuales

				• Porcentaje de No Cobertura: 0.08% • Universo Total: 511 **Proceso de selección y análisis del componente Directores** • Encuesta Probabilista • Tamaño de Muestra Propuesto: N = 121, Muestra Final N1= 111 • Tasa de Respuesta : 91.73% • Error de Muestreo: 8 puntos porcentuales • Porcentaje de No Cobertura: 0.08% • Universo Total: 511 **Proceso de selección y análisis del componente Funcionarios(as)** • Encuesta No probabilística • Muestra recolectada: 143 de 111 Alternativas **Proceso de selección y análisis del componente Nutricionistas y Cocineras** • Encuesta No probabilística • Muestra recolectada: 65 de 111 Alternativas
3.4	**Análisis de datos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali **Se triangulan datos cuanti-cuali pero no de manera generalizada** Se triangulan datos cuanti-cuali de manera generalizada	0 1 **5** 10	Según se apunta en el informe se usó la triangulación de datos al contener algunas de las diferentes técnicas grupos de preguntas y cuestiones comunes. Se observan apartados de los hallazgos donde sí se triangula la información cuantitativa y cualitativa, otros muchos que focalizan únicamente en los datos de la encuesta y otros centrados en datos cualitativos. Es por este motivo que se le otorga una calificación intermedia en esta pregunta.
3.4.2	¿Se cuantifican las respuestas siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	Ninguna / No aplica No se cuantifican las respuestas cuando es posible Se cuantifican las respuestas a veces cuando es posible **Se cuantifican las respuestas de manera generalizada cuando es posible**	0 1 5 **10**	La cuantificación de las respuestas es un recurso que si trata con suficiencia el Informe, especialmente en todos los datos relativos al proceso de encuestación, que constituye la fuente principal de la que se nutre, junto al análisis del marco normativo. También se emplea en ocasiones para las respuestas obtenidas en dinámicas cualitativas, como los talleres, entendiendo que debieron administrar un cuestionario a los participantes para poder establecer dichos porcentajes sobre las respuestas de los participantes.
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica **No se usaron datos del monitoreo** Se usaron datos del monitoreo de manera insuficiente Se usaron datos del monitoreo de manera suficiente	0 **1** 5 10	*Este informe de evaluación no utiliza datos del monitoreo del programa. Son anecdóticos los datos que se presentan, y las contadas ocasiones en que esto se hace se especifican la fuente de la que se tomaron. Son ilustrativos los datos que se han identificado en todo el informe de evaluación:* ▪ *"…El IMAS invirtió en el 2015 un presupuesto de 140.000 millones de colones en el conjunto de proyecto sociales de la institución. De ese monto, un poco más de 18.000.000.00 millones de colones fueron destinados a las alternativas de cuido en el marco de la Red Nacional de Cuido. Este presupuesto ha crecido en los últimos años en forma exponencial así como la cantidad de niños que se atienden. En el 2010 se atendían alrededor de 2.200 niños y en el 2015 unos 18.000…" (Pág. 256)* ▪ *"… Como tercer elemento, se tiene la escasez de recurso humano. El incremento de más de 15,000 niños y niñas en 6 años, con la misma cantidad de recurso humano, ha provocado que ... " (Pág. 265)* ▪ *"… todo es diferente, porque el personal que se dedicaba a este trámite cuando se atendían alrededor de 2000 niños y niñas, sigue siendo el mismo, aun cuando el programa ha alcanzado otras dimensiones y se le proporciona subsidio a más de 18.000 niños tal y como ha sido señalado en el transcurso del documento..." (Pág. 284)* Llama la atención que siendo esta una evaluación de diseño, gestión y ejecución sean estos los datos de ejecución del programa que se expongan en el informe final de evaluación. Los únicos datos que se utilizaron fueron los relativos a la selección de la muestra, requiriendo las bases de datos un amplio trabajo de depuración por parte del equipo evaluador.
3.4.4		Ninguna / No aplica No se recurrió a fuentes externas	0 1	

	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	**Se recurrió a fuentes externas de manera insuficiente** Se recurrió a fuentes externas de manera suficiente	**5** 10	Se usaron fuentes externas, centrando su uso en la contextualización, pero no en el análisis evaluativo propiamente dicho que es lo que se juzga en esta pregunta que se encuentra bajo el subcriterio de metaevaluación precisamente de análisis de datos. Entre los datos empleados en la contextualización se encuentran los siguientes, algunos de ellos ya citados en los propios TdR: Índice de Bienestar de la Niñez y la Adolescencia (IBINA), desarrollado por MIDEPLAN; datos de la Presidencia de la República (2015), Auditoría de Carácter Especial sobre la Red Nacional de Cuido y Desarrollo Infantil (Contraloría General de la República, 2015); VII Estado de los Derechos de la Niñez y la Adolescencia en Costa Rica; estadísticas del PANI; Encuesta Nacional de Hogares del INEC; Censo del 2011 del INEC. Incluso el análisis de esta fuente es complejo ya que no todas aparecen referenciadas en la bibliografía que presenta el Informe de evaluación.
3.5	**Criterios de evaluación**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			El significado de los criterios de evaluación empleados es congruente con los documentos de referencia (MIDEPLAN) y con lo que estipulan los propios TdR. A veces no emplea el criterio de eficiencia al no disponer de datos sobre presupuesto -podrían haber considerado otro tipo de recursos
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica No lo expone Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente **Lo expone de forma concluyente y comprensible**	0 1 5 **10**	La pertinencia es un criterio que se aborda desde diferentes perspectivas en esta evaluación Pertinencia de las AADN en el marco de la Red Nacional de Cuido y Desarrollo Infantil Pertinencia con respecto a las necesidades de la población y las políticas del país Pertinencia con respecto a la Teoría de la intervención de la REDCUDI Pertinencia de la planificación institucional Pertinencia entre los objetivos de las Alternativas, los objetivos institucionales del IMAS y los objetivos del Programa Nacional REDCUDI Análisis de la pertinencia, eficiencia y eficacia de los servicios brindados por las AADN Análisis de la pertinencia, eficacia, eficiencia y sostenibilidad de los procesos relacionados con el subsidio económico Por el amplio peso que tenía este criterio en los TdR, el Informe lo aborda de manera completa.
	b) logró las metas definidas	Ninguna / No aplica No lo expone Lo expone pero no de forma concluyente y comprensible **Lo expone de forma concluyente y comprensible**	0 1 5 **10**	La eficacia es un criterio que se aborda desde diferentes perspectivas en esta evaluación 6.1 Eficacia de los servicios brindados por las AADN 6.2 Eficacia de los procesos relacionados con el subsidio económico. En algún aspecto del análisis las conclusiones presentan afirmaciones no sustentadas suficientemente en hallazgos: Un ejemplo: "El proceso es eficaz en tanto a que, contribuye a alcanzar la población objetivo a la cual la institución se ha comprometido, pues se puede afirmar que los niños y niñas incluidos dentro del proceso subsidio, pertenecen efectivamente a los sectores que se necesita abordar como parte de la política social del país" (Pág.19). Sin embargo no se ha aportado información de seguimiento relativa al perfil de los niños incluidos en el servicio más que la ofrecida por la encuesta. Junto a ello, algunas de las definiciones del criterio eficacia, lo que se juzga es más el respeto del marco normativo vigente que el logro de resultados sobre la población beneficiaria, algo que podría entrar más en la órbita de la pertinencia.
	c) es eficiente	Ninguna / No aplica **No lo expone** Lo expone pero no de forma concluyente y comprensible Lo expone de forma concluyente y comprensible	0 **1** 5 10	Son diferentes ámbitos sobre los que los TdR solicitaban una evaluación de la eficiencia: - En relación con los servicios brindados por las AADN. No se aborda justificando falta de información. - En relación con procesos relacionados con el subsidio económico. En este sentido se valora en qué medida se realiza el proceso con el mejor uso de los recursos y se logran las metas programadas, con el mínimo de recursos disponibles y tiempo, logrando su optimización. A este respecto las fuentes consultadas no eran pertinentes al respecto de ofrecer información relevantes por lo que la fuente de información que permitió alimentar este criterio de evaluación fueron preguntas muy puntuales de la Encuesta realizada a los Directores de las alternativas, algo que se considera muy insuficiente para poder enjuiciar la eficiencia. Pese a ello se identifican factores críticos para optimizar la eficiencia como la estandarización de procesos y el establecimiento de mecanismos de control.

	d) causó impactos (outcome e impacto) deseados y no deseados.	**Ninguna / No aplica**	**0**	No formaba parte del alcance de esta evaluación evaluar el impacto, más allá de su especificación en el marco de la teoría de la intervención que debía reconstruirse, únicas referencias al respecto en los TdR y que sí aborda el Informe.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	No formaba parte del alcance de esta evaluación evaluar la durabilidad del impacto, más allá de su especificación en el marco de la sostenibilidad de los procesos, que si se aborda en el Informe; y la sostenibilidad de la articulación intra e interinstitucional entre los diferentes actores y del sistema de seguimiento y evaluación, que no se abordan en el informe.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	Se hace alguna referencia en el informe a la falta de información disponible para determinados requerimientos de la evaluación, como en el caso del análisis de la eficiencia del subsidio, pero no puede decirse que llegue a exponerse ni que haga de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto). Esto es aún más necesario cuando se indique en la teoría de la intervención que la Secretaria Técnica genera como producto un sistema de información sobre población beneficiaria y Alternativas consolidado, orientando las decisiones y acciones de la REDCUDI. Resulta también relevante dada la ausencia de datos de seguimiento del programa para el mejor análisis de la ejecución del beneficio que se evalúa, así como para caracterizar a la población objetivo y su grado de participación.
		No lo expone	**1**	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	No constituye un foco de la evaluación al no estar dentro de sus objetivos la medición y evaluación de efectos o impactos, ni por tanto de cambios observados. No obstante, en el informe se presentan cambios observados en la población, esencialmente a partir de percepciones cualitativas. Puede entenderse la reconducción de dichos cambios hacia el proyecto como el establecimiento de vínculos claros y directos con el programa en términos de atribución/contribución. En este sentido puede afirmarse que en algunos casos los cambios observados se reconducen al proyecto y se hace de manera concluyente y en otros simplemente se identifica un posible vínculo pero no de manera concluyente, pues los datos en los que se apoyó responden únicamente a percepciones cualitativas no trianguladas con otros datos, por ejemplo de seguimiento del proyecto
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	**5**	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	El Informe no da respuesta a dos de los objetivos de los TdR (5.2.4 y 5.2.5 página 12 TdR). En consecuencia no responde a las dos preguntas que se asociaban a ellos (página 13 TdR): *6.4 ¿Es la estrategia de articulación intra e interinstitucional y su implementación pertinente, eficiente y sostenible?* *6.5 ¿Es el seguimiento y la evaluación de las modalidades de alternativas pertinente, eficiente y sostenible?* *Junto a ello, algunos aspectos considerados como relevantes en los TdR no encuentran respuesta en la evaluación, entre los que destaca el alcance de la evaluación para ofrecer información sobre:* ▪ *Visión de Áreas regionales a partir de muestras representativas, que obtiene una respuesta limitada,* ▪ *Características culturales de la población y climatológicas de la zona en cuanto a las infraestructuras y el equipamiento de las AADN, que no se contempla* ▪ *Modelo de costo de atención según AADN y según características de la zona y de la población que se aborda de forma muy limitada* ▪ *Establecimiento de indicadores para las cadenas de resultados en la reconstrucción de la Teoría de la intervención, que no se aborda.* ▪ *O el producto vinculado con un Documento con líneas estratégicas para el abordaje del plan de acción de las recomendaciones emitidas de la evaluación, que no se incluye*
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	**5**	
		Responde adecuadamente a todas las preguntas	10	
4.2		Ninguna / No aplica	0	
		No se describen puntos fuertes y débiles	1	

	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Se describen principalmente puntos fuertes o débiles **Se describen balanceadamente puntos fuertes y débiles**	5 **10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica No se distinguen con claridad declaraciones descriptivas y valorativas Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe **Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto**	0 1 5 **10**	Es fácil distinguir entre declaraciones descriptivas y valorativas. Constituye un elemento facilitador el hecho de que se presenten de manera diferenciada hallazgos y conclusiones de la evaluación y por lo general las declaraciones valorativas se concentran en el capítulo de conclusiones como corresponde.
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica No se especifica Se especifica en ocasiones **Se especifica en general**	0 1 5 **10**	
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica Las conclusiones no están basadas en los resultados Alguna o algunas conclusiones no están basadas en resultados **Todas las conclusiones están basadas en los resultados**	0 1 5 **10**	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica No se llega ni de forma lógica ni comprensible Se llega de forma lógica o comprensible pero no las dos **Se llega de forma lógica y comprensible**	0 1 5 **10**	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica No son relevantes ni específicas Son relevantes o específicas pero no las dos **Son relevantes y específicas**	0 1 5 **10**	Aunque en general la formulación de las recomendaciones pueden calificarse como específica y muchas de ellas pueden identificarse el destinario de la recomendación, en ocasiones, la formulación no es suficientemente concreta ni especifica qué unidad u órgano debe llevarla a la práctica. Tampoco se acompaña de plan de acción que facilite aclarar la respuesta de gestión en estos casos. A modo de ejemplo, "Se sugiere fortalecer el sistema de planificación mediante una articulación en los ámbitos local, regional y nacional, de manera que incluya una constante actualización de información, que garantice procesos ajustados a la realidad de las áreas regionales del país".
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica El Informe no contiene un resumen El informe contiene un resumen estructurado **El informe contiene un resumen estructurado, bien legible/comprensible**	0 1 5 **10**	El informe presenta un resumen estructurado sobre la base de las preguntas de evaluación
5.	**Exhaustividad y adecuación de la documentación**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica **Los TdR no son accesibles en anexos** Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción Los TdR son accesibles en la lengua común de trabajo o en una traducción	0 **1** 5 10	No se incluyen los TdR en Anexo (Aunque si han formado parte de la documentación de base de esta metaevaluación).

5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	En cuanto a la documentación contenida en los Anexos, el documento que los contiene presenta: **Si se incluye**: Personas entrevistadas y Resultados de cuestionarios agregados **No se incluye**: Cronograma de la fase de campo y listado de lugares
		No se incluyen en anexos	1	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	**5**	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	En cuanto a la documentación contenida en los Anexos, el documento que los contiene presenta: **Si contempla**: Guías para grupos focales y lista de literatura y principal documentación utilizada **No contempla**: Guías para entrevistas y cuestionario para encuesta
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	**5**	
		Se incluyen en anexos (guías y fuentes)	10	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	Ninguna / No aplica	0	El informe presenta una estructura adecuada y que responde en gran medida a la especificada en los TdR (Pág.31), salvo en los puntos que se solicitaban relativos a Lecciones Aprendidas, Propuesta de estrategia para la socialización de resultados, y Propuesta de líneas estratégicas para el abordaje del plan de acción de las recomendaciones. En cuanto a los elementos que facilitan su lectura se contemplan gráficos y tablas pero no resúmenes de los capítulos. Anexo 1. Ejemplos de datos obtenidos en el proceso de consulta Anexo 2. Errores de Muestreo para la Encuesta de Infraestructura y Directores
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	5	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura	**10**	

4.4.6 Evaluación 06: CEN-CINAI

0.1	**Nombre del proyecto / programa evaluado**	**SERVICIO DE ATENCIÓN Y PROTECCIÓN INFANTIL (API) EJECUTADO BAJO LA MODALIDAD DE INTRAMUROS, PERTENECIENTE AL PROGRAMA CEN CINAI (CENTROS DE EDUCACIÓN Y NUTRICIÓN Y CENTROS DE ATENCIÓN INTEGRAL).**
0.2	**Institución que implementa**	Dirección Nacional de Centros de Educación y Nutrición y de Centros Infantiles de Atención Integral (Dirección Nacional CEN-CINAI) del Ministerio de Salud.
0.3	**Evaluadores**	Únicamente en el Resumen Ejecutivo que aparece en la Web figuran los nombres concretos del equipo que ha realizado la evaluación. Sin embargo, no se ha hallado en la documentación consultada referencia alguna a los integrantes concretos del Equipo Gestor, del Equipo de Toma de Decisión o del asesor externo, si bien estas figuras aparecían reflejadas en los TdR. Universidad de Costa Rica (UCR), por medio del Centro de Investigación y Capacitación en Administración Pública (CICAP) y la Escuela de Salud Pública (ESP) Equipo Evaluador: Karol Rojas Araya; Evelyn Zamora Serrano; Cynthia Guzmán Trejos; Ailhyn Bolaños Ulloa; Diana Arce Flores; Esteban Mora Martínez; Víctor Sánchez Brenes. Equipo Directivo UCR: Mayela Cubillo Mora, Directora CICAP; Mauricio Vargas Fuentes, Director ESP
0.4	**Monto del proyecto/ programa**	El único dato económico hallado en el Informe es el siguiente: "El presupuesto total asignado a la Dirección Nacional de CEN CINAI correspondió en el año 2015 a un poco más de 51 mil millones de colones (51.270.426.969,13)". (p. 30)
0.5	**Monto para la evaluación**	No especificado enTdR ni en Informe final de evaluación.

0.6 Descripción del proyecto/programa evaluado
El Programa CEN CINAI es una iniciativa que ha venido desarrollando el Ministerio de Salud de Costa Rica desde la década de 1950, en su búsqueda por contribuir a la mejora del estado nutricional y el desarrollo de la población materno-infantil, que vive en condiciones de pobreza y/o riesgo social; mediante los servicios de nutrición preventiva; atención y protección infantil; y promoción del crecimiento y desarrollo; ofrecidos a escala nacional. En concreto, el Servicio de Atención y Protección Infantil (API), que es el objeto de evaluación, ofrece diariamente a niños y niñas menores de 13 años, de acuerdo a la necesidad de cada familia, los siguientes servicios: a) Servicios de atención integral en un entorno adecuado; b) Promoción del crecimiento y desarrollo infantil; c) Acompañamiento a familias y promoción de servicios comunitarios.
0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL
En los TdR se indica que: - *"Por medio de la orientación metodológica de MIDEPLAN y el Programa FOCEVAL al Ministerio de Salud y puntualmente de la Dirección Nacional de CEN CINAI, se decide evaluar la gestión de los procesos API".* - "*se contará con el apoyo del Programa de Fomento de Capacitación en Evaluación (FOCEVAL) a lo largo de todo el proceso de la evaluación, en donde se aportará la colaboración de un experto internacional en la figura de padrino el cual estaría contribuyendo a la calidad del proceso de evaluación, asesorando tanto al Equipo de Gestión de la evaluación, como al Equipo de Evaluadores que la realicen*". Sin embargo, resulta llamativo que a lo largo del Informe de evaluación no se haga ninguna mención explícita a ese apoyo de FOCEVAL.

0.8 Cualificación cuantitativa (con comentarios) según criterios

72,1 [Sobre 100]	**7,5**	*Imparcialidad e independencia: [Sobre 100]*	A lo largo de la evaluación se va poniendo de manifiesto cómo el equipo externo aporta objetividad y rigor en la adopción de diferentes decisiones referidas a la metodología y el trabajo de campo. Se incorporan múltiples perspectivas, pero fundamentalmente a la hora de recoger la información, no tanto en las fases de análisis o presentación de resultados.
	9	*Equidad, participación y transparencia: [Sobre 100]*	Sobre aspectos de equidad hay poca información (aunque los TdR reflejaban la necesidad de adoptar la perspectiva de género); en cuanto a la participación, estuvo bastante limitada, puesto que muchos aspectos ya venían fijados en el TdR, y la implicación de los actores se tradujo sobre todo como informantes en la recogida de información; en cuanto a transparencia, el balance es positivo, no se observan obstáculos relevantes al respecto.
	33,9	*Credibilidad y exactitud: [Sobre 100]*	Esta evaluación presenta una situación intermedia en este criterio. Hay algunos aspectos del mismo en los que se producen avances importantes (ej. el objeto de la evaluación y los objetivos que se persiguen están claros, pero hay otros aspectos donde el logro no es total (ej. análisis de grupos de interés; utilización efectiva de la teoría de la intervención). En lo que se refiere a diseño y métodos, como aspecto positivo se adopta una pluralidad de métodos, pero por otro lado no siempre hay una suficiente justificación o explicación sobre el porqué de las decisiones metodológicas adoptadas. Los procedimientos de selección presentan cierto margen de mejora, mientras que en el análisis de datos, los resultados son aceptables, y el abordaje de los criterios es bastante satisfactorio.
	17,8	*Calidad del informe y utilidad: [Sobre 100]*	Los aspectos incluidos en este criterio obtienen una alta valoración, la vertebración del informe es visible y comprensible, sigue una secuencia lógica y coherente, los elementos descriptivos conducen ordenadamente a los elementos valorativos, y las recomendaciones resultan útiles y suficientemente concretas. El aspecto que presenta mayor debilidad en este criterio es la escasez de explicaciones o aclaraciones acerca del grado de fiabilidad de los datos obtenidos.
	4	*Exhaustividad y adecuación de la documentación: [Sobre 100]*	Estas cuestiones obtienen una valoración moderada. Los anexos, a pesar de ser extensos y contener múltiple información, no recogen aspectos de especial relevancia, como los TdR o el cronograma y listados de personas o lugares.

0.9 Resumen de fortalezas de la evaluación.

- Se describe de manera detallada el objeto de evaluación y la teoría de intervención reconstruida, lo cual ayuda a la contextualización y aporta riqueza al proceso evaluativo puesto en marcha.
- Enfoque mixto cuantitativo/cualitativo, se concede importancia a la complementariedad entre diversas fuentes y técnicas. En el informe se insertan de forma coherente diversas aportaciones cualitativas literales, que enriquecen la lectura.
- La estructura por criterios, que se visibiliza con claridad y vertebra las conclusiones y recomendaciones, permitiendo así partir de las evidencias y hallazgos, pero alcanzando una visión más global.
- Sistemáticamente, cada hallazgo va acompañado del conjunto de evidencias que lo sustentan, y que han sido obtenidas a lo largo del proceso de evaluación y de la aplicación de las diversas técnicas.
- Se produce una considerable reflexión acerca de las limitaciones, implicaciones y retos, así como las buenas prácticas y lecciones aprendidas en el proceso de evaluación.

0.10 Resumen de debilidades de la evaluación.

- Insuficiente desarrollo y justificación del diseño metodológico adoptado (ej. decisiones, estrategias, técnicas); el Informe entra directa y acríticamente a describir las técnicas utilizadas, sin que se visibilice una reflexión previa sobre su pertinencia para esta evaluación. Tampoco hay evidencias ni se hace explícita en el diseño de la evaluación si se ha incluido, y cómo, la perspectiva de género y de derechos de la niñez y adolescencia
- No hay un abordaje profundo del análisis de datos; fundamentalmente, lo que se señala es que se triangularán los resultados obtenidos por las distintas técnicas aplicadas. Apenas hay alusiones a cómo se ha realizado el análisis cuantitativo (más allá del Anexo con Gráficas y Estadísticas), y en el análisis cualitativo, las estrategias descritas son bastante genéricas.
- No se refleja de forma clara y precisa el número de personas consultadas (total y por cada una de las técnicas).

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	La evaluación fue desarrollada por un equipo multidisciplinar de la Universidad de Costa Rica (UCR), por medio del Centro de Investigación y Capacitación en Administración Pública (CICAP) y la Escuela de Salud Pública (ESP), por encargo del MIDEPLAN. En el resumen ejecutivo que aparece en la Web figuran los integrantes concretos, sin que se haya encontrado información sobre las personas que componían el Equipo Gestor, el Equipo de Toma de Decisión, o el asesor externo.
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	5	
		Evaluación externa (100% externo a instituciones evaluadas)	**10**	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados?	Ninguna / No aplica	0	La pluralidad de perspectivas a partir de las diversas fuentes utilizadas se refleja fundamentalmente en el levantamiento u obtención de información, donde se recurre a diversas técnicas (entrevistas, observación, talleres de consulta participativa, cuestionarios). Se involucró a los diversos niveles de gestión institucional: Central, Regional (direcciones regionales), y Local (oficinas locales y establecimientos CEN y CINAI).
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	**5**	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos)	10	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	***Calificación (numérico)***	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos?	Ninguna / No aplica / No hay información en los documentos consultados	0	No puede decirse que haya existido una involucración activa de los grupos meta o de otros grupos interesados a lo largo del proceso de evaluación. En primer lugar, hay que tener en cuenta que la población objetivo (beneficiaria) son niños y niñas menores de 13 años, cuya implicación en la evaluación resulta más compleja que si se trata de otros perfiles, p.ej. población adulta. No obstante, el Informe recoge algunos aspectos que guardan cierta relación con este asunto: - En el anexo que describe las etapas de la evaluación, en la de diseño de la evaluación, se indica: "*También se realizaron consultas, reuniones y entrevistas no estructuradas con actores particulares relacionados con la prestación del servicio de API, principalmente en el nivel central*". (p. 150). En cualquier caso, no parece que esas acciones supusieran una implicación activa de tales actores, sino más bien que esos perfiles "poseían" información cualitativa y se buscaba recoger esa información. - Por otra parte, en la última etapa, referida a elaboración del informe de evaluación y su socialización, se indica: "*A la entrega del informe borrador le sigue una sesión de devolución de los resultados por medio de una presentación a las autoridades y funcionarios de la Dirección Nacional CEN – CINAI*". (p. 151). Pero de nuevo, parece tratarse de una presentación convencional de resultados, aunque es positivo que se realizara tras la entrega del informe borrador, de forma que las sugerencias u observaciones pudieran tenerse en cuenta de cara a la versión definitiva del informe.
		No hay participación	1	
		Participan escasa o desigualmente en planificación o implementación	**5**	
		Participan activamente en la planificación y en la implementación	10	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	Las preguntas (interrogantes) y las sub-preguntas de evaluación venían fijadas a través de los TdR, y como tal fueron trasladadas al Informe. En líneas generales, presentan claridad, son entendibles, y guardan relación con la evolución y con el contexto en que se desenvuelve el servicio evaluado.
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	5	
		Las preguntas son claras y adaptadas	**10**	
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	Ninguna / No aplica	0	Los TdR presentan una estructura que resulta clara y ordenada, a través de 24 puntos diferenciados: 1.Introducción 2.Antecedentes 3.Alcance ev. 4.Usuarios ev. 7.Fases, plazos, productos 8.Metodología 9.Perfil equipo 10.Presentación de propuestas 13.Plazo,lugar,horaentrega 14.Forma de pago 15.Confidencialidad y protección datos 16.Propuedad intelectual 19.Garantía 20.Aspectos legales 21.Declaraciones y Certificaciones 22.Suscripción y verific cumpl.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	**10**	

				5.Objetivos ev. / 11.Asesor externo / 17.Coordinación y comunicación / 23.Contratista 6.Preguntas ev. / 12.Aprobación productos / 18.Condiciones específicas / 24.Anexos Asimismo, la redacción de cada uno de esos puntos resulta adecuada y comprensible.
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica No hubo un taller de inicio con stakeholders para facilitar su participación No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.) **Hubo un taller de inicio con stakeholders**	0 1 5 **10**	Sí tuvieron lugar talleres iniciales, pero no formaron parte del proceso de evaluación realizado por el equipo evaluador, sino de manera previa, durante la elaboración de los Términos de Referencia y en relación con la construcción de la matriz de evaluación. Así se refleja en los propios TdR: - *"El Equipo Gestor realizó talleres de consulta a actores internos (funcionarios de CEN CINAI) y externos, quienes aportaron insumos para la construcción de esta batería de preguntas (matriz de evaluación)"*. (punto 6, p. 14). Sin embargo, más allá de esto, no se entra en mayor detalle sobre este asunto ni en los TdR ni en el Informe.
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica No son accesibles para el público (ni Informe ni resumen) Son accesibles para el público (Resumen o extracto) **Son accesibles para el público (Informe y resumen)**	0 1 5 **10**	Se entiende por accesibilidad su disposición a través de internet. En este sentido, en Internet es posible acceder al Informe Final, así como al Resumen Ejecutivo. Hay que señalar que el Informe Final que aparece en la Web difiere sensiblemente en algunos apartados del Informe Final facilitado para realizar esta metaevaluación (p.ej. Anexo 8.13). Enlaces web: https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/de78ad1f-d0cc-4eb0-ae13-bf959d0f7a2b/Evaluacion_sobre_gestion_Cen_Cinai.pdf?guest=true (Informe Final). https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/984f6560-5b5e-47f4-8380-dd301ac1b12f/CenCinai_Informe_Final_Resumen_Ejecutivo.pdf?guest=true (Resumen Ejecutivo).
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	Metas, objetos y modelo de cambio	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica No se describe el objetivo de la evaluación (Ni TdR ni Informe) Se describe el objetivo de la evaluación de manera imprecisa **Se describe de manera precisa el objetivo**	0 1 5 **10**	En el apartado 5 de los TdR se reflejan los Objetivos de la evaluación, tanto el general como los cuatro objetivos específicos: Esos mismos objetivos se reflejan en el Informe final de evaluación, dentro del capítulo "Diseño y proceso de la evaluación".
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica No se describe el objeto de evaluación Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes **Se describe con precisión (al menos los ítems contemplados)**	0 1 5 **10**	El capítulo de descripción de la intervención es extenso y ofrece múltiple información, articulándose ese contenido fundamentalmente en torno a la lógica de intervención y la reconstrucción de la cadena de resultados. Asimismo, se ofrece información sobre actores, cuál está siendo el perfil de la población beneficiaria, el presupuesto y los recursos humanos dedicados al servicio, etc. Como elemento de mejora, no obstante, podría señalarse que la información se expusiera de manera más clara y ordenada, ya que en ocasiones se repite información similar en más de una ocasión.
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica No se realizó un análisis de stakeholders **Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.)** Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	0 1 **5** 10	En sentido estricto, lo que se realizó fue una enumeración de los diversos actores cuya intervención es necesaria para que el servicio API pueda realizar su oferta a la población (p. 29 p.ej. funcionarios del Programa; beneficiarios; actores externos…), pero más allá de eso no hay un análisis o profundización en las características de cada actor, sus intereses y necesidades, etc.
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica No se elaboró **Se elaboró pero no se usa en la evaluación o no es concluyente** Se elaboró, se usa en la evaluación y es concluyente	0 1 **5** 10	El capítulo 4 del Informe Final está dedicado a la descripción de la intervención o teoría de la intervención, con una extensión aproximada de 15 páginas. Es decir, se realiza un tratamiento bastante completo de esta cuestión, si bien se echa en falta una estructura algo más clarificadora, para facilitar su lectura. En concreto, el desarrollo de estos aspectos en el Informe se concreta en:

				- Con los insumos generados por la evaluación, se hace una reconstrucción de los objetivos del servicio API, con sus respectivos procesos, metas, indicadores y resultados, reflejados en una tabla. - De una forma más extensa, también se describe el "deber ser" del servicio evaluado a partir de i) la problemática descrita sobre nutrición y desarrollo infantil, ii) la revisión documental, iii) el trabajo de campo, iv) los resultados de la evaluación, y v) los objetivos establecidos por la normativa. - En esta teoría de la intervención específica para el servicio evaluado, se llegan a desglosar los efectos e impactos. También se elabora una cadena de resultados, donde se reflejan los insumos, actividades (procesos), productos/servicios, efectos e impactos (p.31). Sin embargo, posteriormente los hallazgos, conclusiones, recomendaciones, buenas prácticas y lecciones aprendidas no se relacionan de manera directa o explícita con esa teoría de la intervención, sino que en el caso de los hallazgos, se vinculan a los criterios y dentro de estos a los interrogantes (preguntas de evaluación), mientras que las conclusiones y recomendaciones se organizan por criterios. Aun así, se han encontrado en el Informe algunas alusiones que vinculan esa lógica de intervención con el desarrollo del ejercicio evaluativo: - "La evaluación de proceso del Servicio API en relación a la Lógica de la Intervención, permite determinar que este servicio no cuenta con objetivos previamente definidos, más sí con una serie de actividades y de procedimientos, metodologías y abordajes, que hacen que el servicio exista y sea reconocido a nivel institucional". (p. 33) En definitiva, más allá de algunas excepciones, hay que decir que se dedican esfuerzos a reconstruir una teoría de la intervención y cadena de resultados (con información sobre efectos e impactos), lo cual es un elemento valioso, pero que no logra permear o influir totalmente en el conjunto de la evaluación. Seguramente, esa influencia limitada se debe a que son ejercicios de "reconstrucción" por parte del equipo evaluador, es decir, el servicio analizado no opera de facto o expresamente bajo esa estructura y esos supuestos, por ello no resultaría demasiado útil trasladar los resultados de la evaluación a ese esquema. Por ello, se asume, que no siempre es fácil su uso en la evaluación, ya que generalmente las políticas públicas no operan con esa lógica explicita y racionalidad. En todo caso, el establecimiento de indicadores para esa cadena de resultados y tratar de alimentarlos a partir de los datos recabados en la evaluación puede suponer una forma de contraste de dichas teorías que tiene un importante potencial para el aprendizaje. Esto último, supondría el uso de la teoría de la intervención como eje vertebrador de la evaluación, bajo un enfoque de Evaluación Basada en la Teoría (EBT)
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica No se formularon **Se formularon pero no son concluyentes** Se formularon y son concluyentes	0 1 **5** 10	El Informe intenta desarrollar y explicar cómo se producen los impactos en el servicio evaluado, si bien no puede llegar a afirmarse que esas aproximaciones y aportaciones resulten concluyentes o estén totalmente consolidadas. Por una parte, el Informe recoge una "Teoría de los impactos para el servicio API" (p. 26), que señala: - "El Servicio API se basa en la Atención Integral de la Salud como un proceso de construcción social. El grado de contribución está condicionado por actividades de promoción de la salud, prevención de la enfermedad, atención de la morbilidad y de las capacidades disminuidas. - Por ende, el Servicio API persigue que los niños y niñas que desde su gestación hasta la adolescencia, han aprovechado los Servicios API y en su interacción con los determinantes culturales, ambientales y sociales, desarrollen capacidades perceptivas, emocionales, intelectuales y conductuales necesarias para el desenvolvimiento físico y psicológico de su ser como integrante de la sociedad. - Como objetivo superior se considera a la persona como un sujeto de derechos que requieren atención personalizada, que tome en cuenta el contexto familiar y comunitario, tanto en aspectos físicos como

				psicosociales, por lo que, la educación y nutrición temprana son primordiales para su desarrollo e incorporación activa a la sociedad". Por otro lado, se especifican los supuestos respecto a cómo el servicio de API producirá los resultados deseados en la población objetivo. (P.ej. "que se disponga de los recursos materiales y económicos requeridos").
	¿Se diferencia entre los diferentes niveles de impacto?	Ninguna / No aplica No se diferencia entre niveles de impacto **Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)** Se diferencia con claridad los diferentes niveles de impacto	0 1 **5** 10	El Informe consigue diferenciar de manera clara entre efectos e impactos, y ello se refleja en su contenido, p.ej.: - Cadena de resultados (en el gráfico se diferencia entre efectos e impactos). (p. 31). - Tabla 3. Resultados (efectos directos e indirectos) denominados frutos e impactos obtenidos a partir de las sesiones de los Talleres de Consulta Participativos. (Se van enumerando diversos efectos, agrupados según su temática, y se establece una correspondencia con impactos). (p. 32).
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica No se describe el contexto de la evaluación **Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención** Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	0 1 **5** 10	La descripción del contexto de la intervención es más extensa y detallada que la descripción del contexto de la evaluación. De hecho, no se utiliza esa terminología de "contexto de evaluación". En todo caso, en el capítulo dedicado al Diseño y proceso de evaluación, se mencionan una serie de limitaciones, implicaciones y retos para la marcha adecuada de la gestión de la evaluación (También hay un anexo al respecto, con mayor nivel de detalle). Como principales factores de riesgo, se señalan los siguientes: "*…tiempo destinado a fases de trabajo de campos, tiempos de aprobación de productos de consultoría, suministro de información lento por las contrapartes, convocatorias asumidas por el equipo evaluador, uso de técnicas cualitativas que requirieron más logística por reprogramaciones, entre otros. No obstante, estas situaciones no afectaron los objetivos de evaluación, pero si generó un desgaste sobre algunos miembros del equipo, requirió la necesidad de ajustar el presupuesto del proyecto por costos adicionales no estimados y contratar personal de apoyo adicional*". (p. 37 Informe Final). En cualquier caso, el informe no llega a relacionar esos factores con la teoría de intervención del objeto evaluado. También hay que señalar que el capítulo final, sobre Buenas Prácticas y Lecciones Aprendidas, constituye un buen "colofón" al proceso evaluativo realizado, aportando algunas claves para la reflexión sobre cómo se llevan a cabo en la práctica las evaluaciones y el rol de los distintos agentes en ello (ej. equipo evaluador vs. equipo gestor), así como las implicaciones que todo ello tiene en el desarrollo de los trabajos.
3.2	**Diseño y métodos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (Los TdR no explicitan los intereses de la evaluación) El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe) **El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)** Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	0 1 **5** 10	Lo que aquí se ha denominado "intereses explícitos de la evaluación", puede entenderse que equivale a los objetivos y las preguntas y subpreguntas de evaluación, que son aspectos concretados en los TdR y que como tal han sido utilizados en el Informe final. En los TdR, se solicita que el plan de trabajo a realizar por el equipo evaluador incluya el diseño metodológico. Asimismo, se señala que ese equipo "definirá la metodología más adecuada para desarrollar la evaluación según su criterio y experiencia, respetando las condiciones establecidas...". También se indica que la metodología debe implementarse con perspectiva y enfoque de género y de derechos de la niñez y adolescencia, sin embargo no se halla ninguna referencia explícita sobre esto en el Informe. Por su parte, en realidad el Informe no refleja ni visibiliza suficientemente las estrategias y el diseño metodológico adoptado, más allá de relacionar los indicadores de la matriz de evaluación con las técnicas cuantitativas o cualitativas a aplicar. Por tanto, no hay un desarrollo metodológico propiamente dicho, centrándose en la descripción de las técnicas que se van a utilizar. El diseño metodológico se circunscribe y limita a la exposición de la matriz de evaluación, lo cual constituye una visión un tanto limitada del diseño de la evaluación, obviando el detalle de las decisiones metodológicas relevantes adoptadas y sus implicaciones para el proceso o el enfoque evaluativo al que se adscribe, entre otros aspectos. Por todas estas razones, se concede una valoración cuantitativa moderada a este ítem.

<table>
<tr><td rowspan="4">3.2.2</td><td rowspan="4">¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?</td><td>Ninguna / No aplica</td><td>0</td><td rowspan="4">Los propios TdR, además de indicar las preguntas de evaluación a las que se debía dar respuesta, apuntaban, dentro del apartado de “Metodología”, hacia la realización de análisis documental, entrevistas, y consultas a los usuarios de la evaluación.
El proceso de evaluación realizado ha supuesto finalmente la aplicación de diversas técnicas (revisión documental, entrevistas, observación, talleres, cuestionarios), pero no ha habido en el Informe una justificación de por qué resulta conveniente esas técnicas (con ese alcance), en lugar de otras; sino que directamente se describen y cuantifican, sin que medie una reflexión de su oportunidad y pertinencia.</td></tr>
<tr><td>No se explicitan ni justifican el diseño y los métodos aplicados</td><td>1</td></tr>
<tr><td>Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados</td><td>5</td></tr>
<tr><td>Se explicitan y justifican suficientemente el diseño y los métodos aplicados</td><td>10</td></tr>
<tr><td rowspan="4">3.2.3</td><td rowspan="4">¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)</td><td>Ninguna / No aplica</td><td>0</td><td rowspan="4">No hay ninguna mención a análisis contrafactuales, no se explora en ningún momento esa opción metodológica.</td></tr>
<tr><td>No se consideró lo contrafactual</td><td>1</td></tr>
<tr><td>Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró</td><td>5</td></tr>
<tr><td>Se consideró lo contrafactual y se especificó cómo</td><td>10</td></tr>
<tr><td rowspan="4">3.2.4</td><td rowspan="4">¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?</td><td>Ninguna / No aplica</td><td>0</td><td rowspan="4">Las diversas técnicas aplicadas cubren tanto aspectos cuantitativos como cualitativos, si bien se reconoce que en esta evaluación tuvo más preponderancia el enfoque cualitativo, “dado el complejo marco de realidad en que se desenvuelve el Servicio API” (p. 36). Las técnicas utilizadas fueron:
- Cuantitativas: cuestionarios (se pusieron en marcha dos encuestas).
- Cualitativas: entrevistas, talleres de consulta participativos, guía de observación.
- Además, se realizó revisión documental.
Finalmente, indicar que al hablar de los retos de la metodología utilizada (p. 151), se señala esto:
- “El uso de un enfoque de evaluación mixto que utilizara tanto técnicas cuantitativas como cualitativas fue necesario para asegurar la calidad de los resultados de la evaluación. Sin embargo, significó un esfuerzo de trabajo muy grande en tanto que los procedimientos metodológicos cualitativos no se limitaron hasta alcanzar el punto de saturación sino que se prosiguió hasta alcanzar un número definido de visitas a los establecimientos. Ello implicó la realización de un gran número de entrevistas individuales que se pudieron haber realizado de forma más eficiente si se hubieran trabajado como entrevistas grupales o grupos focales”.</td></tr>
<tr><td>Los métodos aplicados no cubren aspectos cuanti ni cuali</td><td>1</td></tr>
<tr><td>Los métodos aplicados cubren aspectos cuanti o cuali</td><td>5</td></tr>
<tr><td>Los métodos aplicados cubren aspectos cuanti y cuali</td><td>10</td></tr>
<tr><td rowspan="4">3.2.5</td><td rowspan="4">¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)?
a) Entrevistas guiadas (número?) ; b) grupos focales (número?)
c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...</td><td>Ninguna / No aplica</td><td>0</td><td rowspan="4">Esta evaluación ha desarrollado diversas técnicas, tanto cuantitativas como cualitativas, habiendo tenido un considerable alcance sobre los diversos perfiles de informantes clave, como se puede ver en la siguiente tabla. No obstante, una de las debilidades de esta evaluación es que no se refleja de forma clara y precisa el número de personas consultadas (total y por cada técnica).
<table>
<tr><th rowspan="2">Técnica / Método</th><th rowspan="2">Tipo</th><th colspan="2">Nº</th></tr>
<tr><th>Ediciones</th><th>Participantes</th></tr>
<tr><td>Entrevistas</td><td>Semiestructuradas en profundidad</td><td>1</td><td>No se indica el número total.
Lo que se dice es:
“Una ASS2 y una TAC por cada uno de los 21 establecimientos seleccionados más la directora en caso de tratarse de un CINAI. Se consideró una (1) ASSC2 y una (1) TAC por establecimiento central y una (1) ASSC2 y una (1) TAC por establecimiento periférico o de zona alejada más 3 ASSC2 y 3 TAC de establecimientos que presentaran alguna innovación!.</td></tr>
<tr><td>Guía de observación</td><td>Presencial</td><td>21</td><td>“Se aplicó en cada uno de los 21 establecimientos visitados”.</td></tr>
<tr><td>Talleres de Consulta Participativa</td><td>Presencial</td><td>9</td><td>160 participantes (Anexo 8.13 Informe Final versión Web). Un taller por cada una de las regiones (9), y se contó con la participación de personal de nivel regional y local.</td></tr>
<tr><td>Cuestionarios</td><td>Telefónica</td><td>1</td><td>224 (112 ASSC2 y 112 TAC)</td></tr>
</table></td></tr>
<tr><td>Se aplicó sólo un método</td><td>1</td></tr>
<tr><td>Se aplicaron dos métodos</td><td>5</td></tr>
<tr><td>Se aplicaron una pluralidad de métodos (más de dos)</td><td>10</td></tr>
</table>

				Cuestionarios | En los Talleres de Consulta Participativa | 1 | No se indica el número total. Personal de Oficina Regional y Oficinas Locales. Revisión documental
3.3	**Procedimiento de selección**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección **Se explicitan en algunos casos (técnicas)** Se explicitan los procedimientos y criterios de selección	0 1 **5** 10	La información sobre este aspecto aparece recogida en uno de los anexos del Informe final, en concreto el de "Técnicas de investigación aplicadas". Se indican cuáles han sido los criterios de selección para algunas técnicas, pero no en todos los casos. Por ejemplo, para algunos perfiles de entrevistas no se indica nada, ni tampoco para los perfiles que fueron invitados a los talleres participativos. Algunos ejemplos: - Entrevistas a personal a cargo del servicio: "El principal criterio de selección para ambas funcionarias, fue que contaran como mínimo con 1 año de laborar en la institución". (p. 148) - Entrevistas a población beneficiaria: "Las madres o padres entrevistados debían contar como mínimo con seis meses de recibir los servicios y fueron escogidos en el momento de hacer la visita al establecimiento de entre quienes llegaban a dejar o recoger al niño o niña". (p.148)
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras?	Ninguna / No aplica La muestra no es adecuada **La muestra no es adecuada o el error no es aceptable** La muestra sí es adecuada y el error aceptable	0 1 **5** 10	La información que se ofrece acerca de las decisiones muestrales es insuficiente. Únicamente se indica que los cuestionarios se seleccionaron a partir de un muestreo aleatorio, y no se indica el margen de error. En el anexo sobre técnicas de investigación aplicadas, dentro del apartado de Cuestionarios, se da alguna información adicional sobre cómo se han seleccionado las unidades últimas a encuestar. (Ej. "*Se aplicaron los cuestionarios a la primera ASSC2 o TAC que contestara el teléfono, reuniera los requisitos y dispusiera del tiempo necesario para realizar la encuesta, en caso contrario se procedía a comunicarse con otra funcionaría ASSC2 y TAC*"). La calificación en esta pregunta no responde tanto a la falta de adecuación de la muestra en sí, sino a la información incompleta que se presenta sobre la encuesta, que limita el rigor del trabajo realizado.
3.4	**Análisis de datos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali Se triangulan datos cuanti-cuali pero no de manera generalizada **Se triangulan datos cuanti-cuali de manera generalizada**	0 1 5 **10**	Según se indica en el capítulo de Diseño y proceso de la evaluación: "Para el análisis de la información, se triangularon los resultados obtenidos por las distintas técnicas aplicadas, de forma tal que permitieran concluir los hallazgos. Resulta pertinente indicar que no todas las interrogantes de evaluación requirieron ser trianguladas debido a que algunas de ellas estaban orientadas a ser contestadas a través de una sola técnica". Puede decirse, por tanto, que esta evaluación tiene presente la triangulación como una herramienta destacada dentro del proceso de análisis de los datos, y que se usa de manera generalizada, con la salvedad de aquellos interrogantes o preguntas de evaluación para las que se contempla recurrir a una única técnica.
3.4.2	¿Se cuantifican las respuestas, siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	Ninguna / No aplica No se cuantifican las respuestas cuando es posible **Se cuantifican las respuestas a veces cuando es posible** Se cuantifican las respuestas de manera generalizada cuando es posible	0 1 **5** 10	Cuando se van describiendo los hallazgos, generalmente se cuantifican las respuestas referidas a la técnica de la encuesta, pero no así las respuestas referidas a otras técnicas de carácter más cualitativo, como las entrevistas o los talleres participativos. Por tanto, esa cuantificación se centra en las técnicas cuantitativas, y no suele afectar a las técnicas cualitativas. Por su parte, el Anexo sobre Gráficas y Estadísticas va presentando información (fundamentalmente gráfica) sobre las distribuciones de frecuencia en las preguntas de las encuestas realizadas, si bien esos datos se muestran solo en porcentaje, sin que aparezcan los datos absolutos.
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica No se usaron datos del monitoreo **Se usaron datos del monitoreo de manera insuficiente** Se usaron datos del monitoreo de manera suficiente	0 1 **5** 10	El hallazgo 13.1 de la evaluación indica que: "No existe un sistema de seguimiento institucional que involucre los diferentes procesos organizacionales, tampoco uno que dé seguimiento a producto o bien a los resultados e impactos de los procesos de CEN CINAI. Al proceso que sí se le da seguimiento periódico es a datos básicos sobre el estado nutricional y de desarrollo de niños y niñas".(P. 69)

				A pesar de eso, en la p.29 del Informe final se aportan algunos datos sobre la población beneficiaria, pero referida únicamente al año 2014 y al conjunto del CEN CINAI, no específicamente al servicio evaluado: - *"…En 2014 CEN CINAI atendió 138.188 clientes y beneficiarias entre bebés, preescolares, escolares, adolescentes madres y mujeres embarazadas o en periodo de lactancia. En total de la población infantil atendida en CEN CINAI, el 8%, de niños y niñas menores de 5 años, presentan desnutrición crónica, según el indicador talla-edad y el 10% de los niños mayores de 5 años presentan delgadez, según el indicador de masa corporal. Mientras que, los resultados de la evaluación del desarrollo se encuentra que el 11% presenta nivel bajo en el área cognoscitiva y 10% en el área del lenguaje"*.
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica	0	En el Informe se refleja el esfuerzo del equipo evaluador por identificar e incorporar información secundaria relevante sobre la problemática vinculada al servicio evaluado (nutrición y desarrollo infantil, pobreza infantil…). De hecho, tanto en la Descripción de la Intervención como en el Anexo "Pobreza y Violencia Infantil", se incluyen fuentes de datos externos, como UNICEF o Estadísticas del Patronato Nacional de la Infancia (PANI). Sin embargo, no puede decirse que esa información se incorpore propiamente a la fase de análisis de la evaluación, sino que son insumos que sirven fundamentalmente para la contextualización de la problemática y del servicio evaluado.
		No se recurrió a fuentes externas	1	
		Se recurrió a fuentes externas de manera insuficiente	**5**	
		Se recurrió a fuentes externas de manera suficiente	10	
3.5	**Criterios de evaluación**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	Uno de los criterios utilizados en esta evaluación es la Pertinencia, y se exponen conclusiones y recomendaciones que se relacionan de forma directa y clara con tales aspectos.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	5	
		Lo expone de forma concluyente y comprensible	**10**	
	b) logró las metas definidas	Ninguna / No aplica	0	Uno de los criterios utilizados en esta evaluación es la Eficacia, y se exponen conclusiones y recomendaciones que se relacionan de forma directa y clara con tales aspectos. Hay que señalar que el criterio de Eficacia se presenta conjuntamente con el de Eficiencia (Eficiencia y eficacia).
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	c) es eficiente	Ninguna / No aplica	0	Uno de los criterios utilizados en esta evaluación es la Eficiencia, y se exponen conclusiones y recomendaciones que se relacionan de forma directa y clara con tales aspectos. Hay que señalar que el criterio de Eficiencia se presenta conjuntamente con otros, según proceda. (Pertinencia y eficiencia, o Eficiencia y eficacia).
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	d) causó impactos (outcome e impacto) deseados y no deseados.	**Ninguna / No aplica**	**0**	Dentro de los objetivos de esta evaluación no se contempla la medición y evaluación de efectos o impactos, si bien ese tema sí se aborda desde el punto de vista del diseño, cuando se reconstruye en el capítulo 4 la lógica de intervención y la cadena de resultados.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	Dentro de los objetivos de esta evaluación no se contempla la medición y evaluación de efectos o impactos, si bien ese tema sí se aborda desde el punto de vista del diseño, cuando se reconstruye en el capítulo 4 la lógica de intervención y la cadena de resultados.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue	Ninguna / No aplica	0	Dentro de las conclusiones sobre Eficiencia y Eficacia (p. 87), se señala la ausencia de sistemas de seguimiento o de información de carácter oficial, sobre el servicio evaluado:
		No lo expone	1	

	usado para la gestión (calidad de la gestión del proyecto)	Lo expone pero no de forma concluyente y comprensible	5	- "La toma de decisiones que se realiza en el Programa no está respaldada por un sistema de información institucional que sea práctico, confiable, seguro y oportuno. Esto limita el seguimiento y la medición de resultados, efectos y de impactos que retroalimenten a la gestión a nivel de aciertos, desaciertos y cambios requeridos en los procesos". - "La ausencia de un sistema de seguimiento institucional, enlazado a un sistema de información, también inexistente, pone en peligro la recolección y análisis de datos confiables, sobre actividades e intervenciones sustantivas como el seguimiento del crecimiento y desarrollo de la población atendida. Esto también compromete la producción de evaluaciones de impacto de los procesos sustantivos, que pueden llevar a la toma de decisiones". No cabe por tanto una utilización generalizada de estos sistemas en la gestión del servicio evaluado, ya que apenas están conformados.
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	**Ninguna / No aplica**	**0**	No procede, puesto que dentro de los objetivos de esta evaluación no se contempla la medición y evaluación de efectos o impactos, ni por tanto de cambios observados.
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	5	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	La forma de estructurar los hallazgos de evaluación incluidos en el Informe, y las evidencias en que éstos se sustentan, se corresponde totalmente con las 24 preguntas de evaluación listadas en el TdR ("interrogantes"), y que están incluidas en la matriz de evaluación. Por su parte, las conclusiones del Informe se estructuran en un plano más global, a través de los diversos criterios contemplados, distinguiendo entre conclusiones generales y específicas para cada uno de ellos. Este mismo esquema se utiliza para las recomendaciones.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	El informe presenta un equilibrio entre la exposición de puntos fuertes y puntos débiles referidos al servicio evaluado. De hecho, en el resumen ejecutivo, los diversos hallazgos y conclusiones se catalogan en ocasiones como "oportunidad", "fortaleza" o "debilidad". Ejemplo de punto fuerte: "*El diseño y la ejecución de los procesos técnicos son pertinentes con la problemática del entorno, con el ordenamiento jurídico y sus correspondientes objetivos del servicio, normas y procedimientos*". (p. 84). Ejemplo de punto débil: "*La supervisión y el acompañamiento no se están brindando con la periodicidad necesaria para la ejecución con calidad de los procesos de API*". (p. 85).
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	El informe refleja de forma adecuada y diferenciada esos dos planos, tanto el descriptivo como el valorativo. El capítulo de hallazgos presenta un tono descriptivo y neutro, se busca el rigor acompañando cada hallazgo de las evidencias en las que se sustenta (epígrafe de hallazgos y epígrafe de evidencias). Por su parte, en el capítulo de conclusiones y recomendaciones ya se introducen de manera notoria los elementos valorativos / evaluativos, a partir de las evidencias y hallazgos encontrados.
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	**10**	
4.4	¿Se especifica cómo las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	El anexo que explica las "limitaciones, implicaciones y retos", podría ser el lugar del Informe donde aparecieran este tipo de consideraciones, pero no hay ninguna mención al respecto. Más bien, se habla de aspectos como la implementación del trabajo de campo y las cuestiones logísticas que de ello se derivó. Hay otro apartado del Informe, en concreto el que lo cierra antes de comenzar los anexos, y que se denomina "Buenas prácticas y lecciones aprendidas" (p. 101 y ss.), donde aparecen enseñanzas muy relevantes derivadas del proceso de evaluación realizado. Por su pertinencia para la cuestión aquí tratada, se rescata lo siguiente, que trata sobre la posible introducción de sesgos por interferencia del equipo gestor en la labor evaluativa:
		No se especifica	**1**	
		Se especifica en ocasiones	5	
		Se especifica en general	10	

				- "El equipo evaluador definió los instrumentos a utilizar y la organización del trabajo de campo. Esto se logró luego de que insistiera en que no era pertinente que el equipo gestor conociera o influyera en la construcción de instrumentos y en organización del trabajo de campo. Sin embargo, siempre se encontraron interferencias, como por ejemplo: que al llegar a algún Centro, ya se tuviera definido con anterioridad cuál TAC, ASSC2 y beneficiario(a) iba a responder a la entrevista. Por lo que el equipo evaluador tuvo que adoptar medidas para asegurar la objetividad del proceso". Esta evaluación aborda la confiabilidad de la información, pero no tanto como aspecto dentro del proceso evaluativo, sino como una de las cuestiones que se quiere conocer acerca del servicio evaluado, por ejemplo en relación con sus sistemas de seguimiento y de información. (Ej. *"Hallazgo 12.1. El no contar con un sistema de información seguro y confiable, genera desconfianza entre los funcionarios, sobre el destino de la información que ellos generan*"). Finalmente, señalar que dentro del Informe se encontró esta alusión que puede guardar relación con los aspectos aquí tratados, pero que se presenta de forma aislada, en una única ocasión y no impregna el conjunto del informe: - (Análisis de información cualitativa "Se le dio credibilidad a los datos mediante el criterio de rigor de confirmabilidad, es decir, en el que aunque no se deseó ocultar la subjetividad del investigador, sí se buscó que los datos y las conclusiones fueran confirmadas por voces de los/as entrevistados". (p. 37).
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica Las conclusiones no están basadas en los resultados Alguna o algunas conclusiones no están basadas en resultados **Todas las conclusiones están basadas en los resultados**	0 1 5 **10**	Las conclusiones se presentan en un capítulo diferenciado al de los hallazgos o resultados, y están estructuradas por criterios (pertinencia; pertinencia y eficiencia; eficiencia y eficacia; sostenibilidad; calidad), desglosándose en cada uno de ellos en una conclusión general y en varias conclusiones más específicas. Esto permite una lectura más global, y a su vez es posible recuperar la trazabilidad de esas conclusiones con los hallazgos a través de la Tabla 5 (p. 90-91). En definitiva, las conclusiones guardan una coherencia y correspondencia directa con los hallazgos o resultados de la evaluación.
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica No se llega ni de forma lógica ni comprensible Se llega de forma lógica o comprensible pero no las dos **Se llega de forma lógica y comprensible**	0 1 5 **10**	La forma de redactar las recomendaciones es igual a como se realiza en las conclusiones: organizadas por criterios, y dentro de cada uno, general y específicas. Además, la Tabla 6 (p. 100-101) refleja de forma explícita el resumen de recomendaciones y su relación con las conclusiones, por tanto la trazabilidad está asegurada, y se llega de forma lógica y comprensible de las recomendaciones a las conclusiones.
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica No son relevantes ni específicas Son relevantes o específicas pero no las dos **Son relevantes y específicas**	0 1 5 **10**	Se dedican aproximadamente 10 páginas del Informe final a las recomendaciones, y éstas se desglosan en recomendaciones de carácter general y de carácter específico, agrupadas según los criterios que han guiado la evaluación. De su lectura puede concluirse que son relevantes y tienen concreción, si bien se echa en falta indicar expresamente sobre qué actor recae la responsabilidad de implementación o de seguimiento de cada recomendación. En este sentido, en el propio Informe se realiza la siguiente aclaración: - "No se operacionalizan las recomendaciones según el destinatario y el horizonte temporal debido a que no se considera adecuado que el equipo evaluador interfiera con ello en la dinámica interna o en la toma de decisiones institucionales." (p. 91).
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica El Informe no contiene un resumen El informe contiene un resumen estructurado **El informe contiene un resumen estructurado, bien legible/comprensible**	0 1 5 **10**	En términos generales, el resumen ejecutivo incluido al principio del Informe Final contiene información suficiente que refleja en qué ha consistido el proceso de evaluación, y qué hallazgos, conclusiones, recomendaciones y buenas prácticas/lecciones aprendidas se han obtenido.
5.	**Exhaustividad y adecuación de la documentación**			
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica **Los TdR no son accesibles en anexos** Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción Los TdR son accesibles en la lengua común de trabajo o en una traducción	0 **1** 5 10	Los TdR no aparecen en el Informe Final, ni tampoco en ninguno de sus anexos.

5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	No se incluye el cronograma de la fase de campo, si bien en los anexos se describen las diferentes etapas del proceso de evaluación, y se dan algunos detalles referidos al trabajo de campo. En el anexo de "Técnicas de investigación aplicadas" se describen los perfiles entrevistados, pero sin especificar las personas concretas. En cuanto a los talleres participativos, se indica que se realizó un taller por cada una de las regiones, siendo un total de 9. Únicamente en el Informe Final que aparece en la Web (que es ligeramente distinto al facilitado para realizar esta metaevaluación), en el último Anexo (8.13), se detalla región, fecha de realización, cantidad de asistentes, y puestos. En definitiva, en esta evaluación se ofrece información sobre el trabajo de campo y las técnicas empleadas, pero no alcanza tanto nivel de detalle, lo más habitual es que se limite a describir el proceso en general, o los perfiles que participaron.
		No se incluyen en anexos	1	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	**5**	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	El anexo 8.12 se refiere a: "Revisión documental: Lista de documentos e información recibida por parte de CEN-CINAI". Por otra parte, hay un anexo denominado "Gráficas y Estadísticas", que, si bien no reproduce como tal los cuestionarios empleados, va reflejando los resultados para cada uno de los aspectos tratados en la encuesta. Sin embargo, no se incluye en el Informe ni en los anexos ningún tipo de guion sobre las herramientas cualitativas (entrevistas, talleres, observación).
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	**5**	
		Se incluyen en anexos (guías y fuentes)	10	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos?	Ninguna / No aplica	0	En términos generales, el informe sigue una secuencia lógica en la exposición de sus contenidos, y su lectura es comprensible. En todo caso, en algunos momentos se observa una cierta "densidad" en los textos aportados, que podrían haber sido aligerados o complementados con algún recurso como una tabla o un gráfico. A pesar de eso, hay que decir que a lo largo del Informe, hay un aceptable número de tablas, no ocurre así con los gráficos, que se presentan únicamente y de manera concentrada en el anexo "Gráficas y Estadísticas". En algunos capítulos, como el de Descripción de la intervención, hubiera sido recomendable incluir algunas explicaciones más explícitas sobre cómo se iba estructurando la información, ya que a veces resultaba difícil para el lector seguir el hilo argumental, se producían "saltos" o se repetía alguna información. Otra debilidad es que los anexos no siguen un orden totalmente lógico, por ejemplo aspectos esenciales como la matriz o las etapas de la evaluación no aparecen de los primeros anexos, sino que vienen precedidos por otros anexos que tratan cuestiones más específicas o puntuales. Por último, indicar que no se incluyen en ningún caso resúmenes de los capítulos.
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	**5**	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	10	

4.4.7 Evaluación 07: EMPLÉATE

0.1	**Nombre del proyecto / programa evaluado**	**PROGRAMA EMPLÉATE**
0.2	**Institución que implementa**	Dirección Nacional de Empleo del Ministerio de Trabajo y Seguridad Social (MTSS).
0.3	**Evaluadores**	Equipo Evaluador: Ericka Valerio Mena; Javier Alfaro Araya; Mª Elena Fonseca Quirós; Mario González Zúñiga; Tabatha Carvajal Ruiz. Equipo Tomador de Decisiones: Amparo Pacheco Oreamuno; Andrés Romero Rodríguez; Florita Azofeifa Monge;Sandra Mongalo Chan; Tatiana Ramírez Chavarría. Supervisado por: Victoria Sánchez Esteban (Consultora Internacional) y Gabriela León Segura (Consultora Nacional).
0.4	**Monto del proyecto/ programa**	*Según consta en Cuadro 2, p. 36:* *Recursos asignados: ¢20.440 millones.* *Recursos ejecutados: ¢15.483 millones. Es decir, un 75,7% de los recursos asignados.*
0.5	**Monto para la evaluación**	*No especificado en Informe final de evaluación*

0.6 Descripción del proyecto/programa evaluado

El Programa Empléate, creado en 2010 y que se integra dentro del Programa Nacional de Empleo (MTSS), constituye una estrategia público-privada de promoción del empleo para personas jóvenes (17-24 años, ampliable a 35 años en caso de discapacidad) en situación de vulnerabilidad (que no estudian ni trabajan, y se ubican por debajo de la línea de pobreza). Esta estrategia se fundamenta en alianzas entre instituciones públicas, sector empresarial y organizaciones no gubernamentales con el propósito de capacitar a esos jóvenes para su inserción al mercado laboral, fortaleciendo su empleabilidad a través de capacitaciones técnicas/ ocupacionales vinculadas con la demanda ocupacional insatisfecha, partiendo de los señalamientos del mercado laboral. A través de este programa se les brinda el apoyo económico necesario, la asesoría y orientación, y la posibilidad de acceder a un proceso formativo.

0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL

Para la elaboración de esta evaluación se conformaron dos equipos, uno Tomador de Decisiones, y otro equipo Evaluador integrado por representantes del Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) y del Ministerio de Trabajo y Seguridad Social) (MTSS).

El proceso evaluativo se desarrolló bajo la modalidad de *Aprender Haciendo*, con el fin de generar capacidades y conocimientos en el tema de la evaluación, replicables a lo interno de las instituciones. Se contó con el apoyo del Programa de Fomento de las Capacidades en Evaluación en países seleccionados en América Latina (FOCEVAL) como órgano financiador, el cual realizó la contratación de dos consultoras, una internacional y otra nacional, responsables de orientar y asesorar al equipo evaluador.

En el apartado 3.4., sobre limitaciones del proceso de evaluación, se incluyen valoraciones sobre cómo resultó la ejecución de la evaluación bajo esta modalidad de Aprender Haciendo (p.44), subrayándose como principales limitaciones las siguientes:

- Diferentes niveles de experiencia en los integrantes del equipo (respecto a su participación en evaluaciones, así como a la aplicación de técnicas de recolección y análisis de la información). Puesto que esta modalidad de *Aprender Haciendo* exige que todos los integrantes participen en todas las actividades del proceso, independientemente de su experiencia, eso conduce a una ralentización, dedicando más tiempo a cada una de las etapas de la evaluación.
- Se produjo rotación de personal dentro del equipo evaluador, que obligó a generar procesos de inducción a funcionarios que ingresaban en sustitución de otros.
- Los integrantes del equipo evaluador tenían en realidad un tiempo limitado para dedicar a la evaluación, ya que generalmente su cargo conllevaba también cumplir con otras actividades.
- Se consideró insuficiente el acompañamiento por parte de las consultorías; los evaluadores sintieron la necesidad de contar con un apoyo técnico permanente.

0.8 Cualificación cuantitativa (con comentarios) según criterios

71,7 [Sobre 100]	**7,5**	***Imparcialidad e independencia:*** [Sobre 10]	El hecho de que la propia institución evaluada participara en la realización de la evaluación pudo ser un factor limitante, sin embargo esto se contrarresta con la utilización de una pluralidad de técnicas y fuentes, que ayuda a tener una visión global y completa.
	5,5	***Equidad, participación y transparencia:*** [Sobre 10]	Esta evaluación no destaca especialmente por la implicación y participación de los diversos actores implicados, que son involucrados generalmente desde el rol convencional de suministradores de información (informantes clave). No obstante, en cuanto a transparencia los resultados son satisfactorios, ya que los productos de la evaluación pueden ser consultados por cualquier persona a través de Internet, y las preguntas de evaluación son claras y permiten hacerse una idea adecuada de los focos donde se centra esta evaluación.

	34,7	***Credibilidad y exactitud:*** [Sobre 50]	Las puntuaciones en las que se desglosa este criterio resultan desiguales en esta evaluación. Hay aspectos que están consolidados, como la descripción del objeto de evaluación; la elaboración y utilización de una teoría de la intervención y cadena de resultados; la utilización de métodos mixtos, que cubren tanto lo cuantitativo como lo cualitativo, así como muestras y errores aceptables; también se saca algún provecho de los datos del monitoreo; y se trabaja desde una óptica de criterios. Como aspectos a mejorar, los objetivos de la evaluación son un tanto genéricos, y se echa en falta el desarrollo de un diseño o modelo de evaluación, puesto que lo que se plasma son fundamentalmente las preguntas de evaluación y los criterios, y las técnicas o fuentes con las que se le da respuesta. Tampoco hay apenas atención al análisis de actores, y todo lo referido a análisis de impactos se plantea como propuesta metodológica a futuro.
	17,5	***Calidad del informe y utilidad:*** [Sobre 20]	Aquí también las puntuaciones en las que se desglosa este criterio resultan desiguales. Como puntos positivos, el Informe da respuesta a las preguntas de evaluación, sigue una secuencia lógica y coherente, presenta un resumen adecuado, y no se entremezcla de forma inadecuada lo descriptivo con lo valorativo. No obstante, como aspectos que se podrían haber enfocado mejor, el informe tiene cierta tendencia a resaltar las debilidades más que las fortalezas del objeto evaluado, y tampoco hay una excesiva atención a las cuestiones de fiabilidad en los hallazgos o conclusiones presentados.
	6,5	***Exhaustividad y adecuación de la documentación:*** [Sobre 10]	Por una parte, hay distintos elementos en el informe que facilitan su lectura, y los anexos contienen información valiosa, si bien se detectan algunas carencias (ausencia de TdR, cronograma de trabajo de campo, o listados detallados de participantes y lugares).

0.9 Resumen de fortalezas de la evaluación.

Esta evaluación se orientó a valorar los niveles de diseño, proceso y resultados del Programa Empléate, con el objetivo de apoyar la toma de decisiones para la mejora del mismo. El período abarca los años 2012-2015. En ella se buscó identificar fortalezas y debilidades de la intervención, con el fin de proponer medidas correctivas para mejorar y facilitar la implementación del Programa. Así, el proceso de Evaluación Empléate cumple con varios propósitos: aprendizaje, mejora, transparencia y rendición de cuentas.

Como fortalezas de la evaluación, destacan:

- *En general, estructura clara, ordenada y coherente, y se incluye información suficiente sobre el objeto de evaluación y sobre el proceso de evaluación desarrollado.*
- *Enfoque mixto, con técnicas cuantitativas y cualitativas. Diversidad de herramientas (tres encuestas, entrevistas, grupos focales, observación, revisión documental y de expedientes).*
- *Abundante información, y de contenido relevante. El contenido y la forma en que están redactadas las recomendaciones contribuye positivamente a su puesta en práctica.*
- *Visión detallada y crítica, aunque eso conduce a que tengan un mayor peso la descripción/valoración de debilidades del Programa que sus fortalezas.*

0.10 Resumen de debilidades de la evaluación.

Como debilidades de esta evaluación, cabe subrayar:

- *En base a la matriz diseñada, la evaluación transita directamente desde las preguntas de evaluación a las técnicas utilizadas, sin que se refleje modelo teórico o metodológico alguno que guíe y justifique las diversas decisiones metodológicas adoptadas.*
- *Escaso protagonismo de los criterios a la hora de estructurar los hallazgos, conclusiones y recomendaciones de la evaluación. Se opta por guiar la estructura según las tres dimensiones de diseño, procesos y resultados.*
- *Ausencia de un análisis previo de stakeholders; tampoco se celebra ningún taller de inicio con agentes involucrados. Se asume la situación de partida como algo fijo, seguramente porque hay otros aspectos que requieren mayor atención e interés por parte de la evaluación, y también por las dificultades para reunir o involucrar a esos agentes. Al hilo de lo anterior, el hecho de que no se contara con la participación (activa y necesaria) del responsable del Programa Empléate, es una limitación que quita fortaleza a la evaluación.*
- *Aunque los Anexos tienen una extensión considerable dentro del Informe, y presentan información diversa y de utilidad, no van precedidos de un índice ordenado y numerado que ayude a identificar el contenido de cada uno de ellos, lo cual resta calidad y puede suponer un menor aprovechamiento de los mismos por parte de quien realiza la lectura.*
- *A nivel de formato, hubiera sido deseable una estructuración algo más precisa y clarificada de los diferentes epígrafes y subniveles que componen cada capítulo.*

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica Evaluación interna (100% trabaja en instituciones evaluadas) **Evaluación mixta (externos y trabajan en instituciones evaluadas)** Evaluación externa (100% externo a instituciones evaluadas)	0 1 **5** 10	La evaluación fue realizada por un equipo de funcionarios del Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) y del Ministerio de Trabajo y Seguridad Social (MTSS) destacados en la Dirección General de Desarrollo Social y Asignaciones Familiares (DESAF) y en la Dirección General de Planificación del Trabajo (DGPT). Además, se contó con una asesora nacional y otra internacional, ambas expertas en materia de empleo. En esta evaluación se utilizó la metodología de "Aprender Haciendo", de modo que se brindan conocimientos teóricos sobre evaluación combinados con una aplicación práctica a través de esta metodología. El equipo que ejecuta la evaluación (equipo evaluador) no está compuesto por evaluadores profesionales, sino personas que desean adquirir conocimientos y experiencia en evaluación, para posteriormente aplicarlos en sus diferentes áreas de trabajo.
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y el análisis de datos, y en la presentación de los resultados	Ninguna / No aplica Las fuentes elegidas no permiten capturar una pluralidad de perspectivas Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas **Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas**	0 1 5 **10**	En esta evaluación se tuvieron en cuenta fuentes primarias y secundarias de información, y se aplicaron tanto técnicas cualitativas como cuantitativas en la recolección de información. En concreto: - Entrevistas a diversos perfiles (18 entrevistas). - Encuesta a diversos perfiles (33 gestores de empleo; 30 centros de formación; 950 beneficiarios). - Grupos focales (8 grupos: 3 gestores de empleo, 3 centros de formación, 2 beneficiarios). - Observación a través de Ventanilla de Atención y Reto. - Revisión documental (3.000 expedientes que aportaron datos sobre 2.822 beneficiarios, así como documentos varios) y bases de datos (PRONAE, DESAF). Por tanto, las fuentes elegidas contemplan una pluralidad de perspectivas, e implican a los principales informantes clave. No obstante, tales fuentes estuvieron sometidas a diversas limitaciones, que quedan expresamente reconocidas en el Informe final de evaluación, p. ej. escasa información documentada sobre actividades del programa; insuficiente información de seguimiento; o dificultades para la aplicación de encuestas.
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	***Calificación (numérico)***	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos?	Ninguna / No aplica / No hay información en los documentos consultados **No hay participación** Participan escasa o desigualmente en planificación o implementación Participan activamente en la planificación y en la implementación	0 **1** 5 10	El informe no refleja la involucración activa de los grupos meta u otros grupos interesados. Se desconoce si hubo intención expresa de lograr esa implicación, si bien el informe menciona expresamente algunas dificultades o limitaciones que pudieron afectar también a esta cuestión: - *Ausencia del responsable del Programa en el proceso de evaluación*, aunque en ocasiones fue representado por otros funcionarios de la unidad ejecutora. Como consecuencia de esta ausencia, hubo falta de retroalimentación y validación por su parte acerca de los avances de la evaluación, lo cual puede poner en riesgo la utilización de los resultados de la evaluación. - *Dificultades para contactar, ubicar e implicar a los beneficiarios del Programa en el proceso de evaluación*, por su situación económica vulnerable, los cambios de residencia, las obligaciones familiares o de otro tipo, etc. No obstante, y aunque sea únicamente a modo descriptivo, el informe sí recopila los diversos agentes involucrados en el Programa (Tabla 2).
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica Las preguntas ni son claras ni adaptadas Las preguntas o son claras o adaptadas **Las preguntas son claras y adaptadas**	0 1 5 **10**	Las preguntas están adaptadas al contexto y pueden considerarse claras. Dichas preguntas se exponen de forma conjunta y sintética en una tabla, vinculando cada una de ellas con la dimensión a la que pertenece (diseño, proceso, resultados), y con el criterio o criterios de evaluación que subyace en cada una de ellas. En total se plantean 13 preguntas de evaluación, correspondiendo casi la mitad de ellas (6) a la dimensión de procesos, lo cual, aunque en principio no tiene por qué resultar problemático, puede ser un indicio de que la evaluación acaba estando más orientada a dar respuesta a aspectos de los procesos, en comparación con las otras dos dimensiones también analizadas (diseño, con 3 preguntas, y resultados, con 4). En general, todas las preguntas de evaluación planteadas tienen un alcance similar, sin que pueda decirse que haya una disparidad de alcance entre ellas, de forma que no resultan ni demasiado específicas ni demasiado generales. La primera pregunta englobada en Resultados ("¿Existe igualdad de acceso a Empléate en cada una de sus fases"?), en realidad podría considerarse más bien una pregunta referida a cuestiones de diseño, si bien seguramente se incluyó ahí por su repercusión a nivel de resultados.

2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica**	**0**	No se dispone de los TdR, si bien las preguntas de evaluación, que son un elemento central en la articulación de la evaluación, resultan suficientemente claras y comprensibles.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	El informe no recoge la celebración de ningún taller de inicio con stakeholders para facilitar su participación. Se desconoce si hubo intención de realizar ese tipo de taller, si bien el informe menciona expresamente algunas dificultades o limitaciones que pudieron afectar también a esta cuestión: - *Ausencia del responsable del Programa en el proceso de evaluación*, aunque en ocasiones fue representado por otros funcionarios de la unidad ejecutora. Como consecuencia de esta ausencia, hubo falta de retroalimentación y validación por su parte acerca de los avances de la evaluación, lo cual puede poner en riesgo la utilización de los resultados de la evaluación. - *Dificultades para contactar, ubicar e implicar a los beneficiarios del Programa en el proceso de evaluación*, por su situación económica vulnerable, los cambios de residencia, las obligaciones familiares o de otro tipo, etc.
		No hubo un taller de inicio con stakeholders para facilitar su participación	**1**	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	
		Hubo un taller de inicio con stakeholders	10	
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica	0	Se entiende por accesibilidad su disposición a través de internet. En este sentido, en Internet es posible acceder al informe completo de esta evaluación, donde se incluye el resumen ejecutivo y también los Anexos. Enlace web: https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/eb1e936e-b470-482b-a9cf-e880c1cff452/Evaluacion_de_dise-o_proceso_y_resultados_del_Programa_Empleate.pdf?guest=true
		No son accesibles para el público (ni Informe ni resumen)	1	
		Son accesibles para el público (Resumen o extracto)	5	
		Son accesibles para el público (Informe y resumen)	**10**	
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	Los objetivos de la evaluación quedan indicados en la p. 40 del Informe final: Objetivo General: apoyar la toma de decisiones para la mejora del Programa Empléate. Objetivos Específicos: - *Diseño*: identificar las posibilidades de mejora del Programa Empléate de cara a las necesidades de la población beneficiaria. - *Procesos*: valorar los procesos de gestión para mejorar la prestación del servicio. - *Resultados*: conocer los resultados del Programa en términos de logro de los objetivos planteados. Asimismo, queda perfilado cuál sería el objetivo de una futura evaluación de impacto. Los objetivos son un poco genéricos, y presentan poca concreción. Resulta deseable una mayor concreción de los objetivos, resaltando en cada uno de ellos aquella cuestión o dimensión sobre la que se desea poner énfasis, mostrando aquellos aspectos que más preocupan sobre la intervención evaluada o sobre los que existe más incertidumbre.
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	
		Se describe el objetivo de la evaluación de manera imprecisa	**5**	
		Se describe de manera precisa el objetivo	10	
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica	0	El Informe final dedica 15 páginas a la descripción del Objeto de Evaluación, deteniéndose en la descripción y explicación de múltiples aspectos sobre ese objeto, a saber: - Problemática que da origen a la intervención. - Intervención, y ubicación territorial del Programa. - Modalidades del Programa. - Focalización de las acciones del Programa. - Procesos de intervención. - Agentes involucrados en el Programa. - Ejecución programática y financiera.
		No se describe el objeto de evaluación	1	
		Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes	5	
		Se describe con precisión (al menos los ítems contemplados)	**10**	

				- Cadena de resultados. En definitiva, el Informe recoge información suficiente y relevante sobre el objeto de evaluación, lo cual debe ser puesto en valor especialmente porque dentro de las limitaciones en el proceso de evaluación aparecen algunas que afectan directamente a la delimitación y conocimiento de ese objeto, como: escasa información documentada sobre actividades del programa; ausencia del responsable del Programa durante el proceso de evaluación.
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica **No se realizó un análisis de stakeholders** Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.) Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	0 **1** 5 10	De la lectura del informe de evaluación, se deduce que había suficiente conocimiento previo acerca de los agentes implicados en el programa, y que la contactación con algunos de ellos (ej. beneficiarios) resultaba difícil, razones todas ellas que pueden haber motivado que no se realizara de forma expresa un análisis de esos stakeholders enfocado a esta evaluación.
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica No se elaboró Se elaboró pero no se usa en la evaluación o no es concluyente **Se elaboró, se usa en la evaluación y es concluyente**	0 1 5 **10**	Una de las preguntas de evaluación, incluida dentro de la dimensión Diseño, alude a: "¿La cadena de resultados del Programa está adecuadamente definida?". Consecuentemente, en los hallazgos de evaluación hay un epígrafe reservado a esta cuestión, y además de eso, y de manera previa, en la descripción del objeto de evaluación queda definida esa cadena de resultados, que tuvo que ser elaborada por el equipo evaluador, ya que no existía de forma explícita y era necesaria para conocer el funcionamiento del Programa y sus resultados esperados. Esa cadena "teórica" (Anexo 25) fue sumamente relevante, ya que se utilizó como insumo para la confección de la matriz de evaluación. Tras la aplicación de las diversas técnicas y la realización de los análisis correspondientes, esa cadena se adaptó a la realidad efectiva del Programa (variando fundamentalmente aspectos relacionados con los efectos e impactos), y se simplificó para que su manejo fuera más sencillo (Figura 3, p. 38). En las conclusiones y recomendaciones no aparecen menciones expresas a la cadena de resultados, aunque esa visión global que se construyó sobre el Programa subyace en el conjunto de la evaluación. Puede decirse, por tanto, que se elaboró la cadena de resultados, se utilizó, y que resultó concluyente.
	¿Se formularon hipótesis de impacto concluyentes?	**Ninguna / No aplica** No se formularon Se formularon pero no son concluyentes Se formularon y son concluyentes	**0** 1 5 10	Esta evaluación establece un objetivo específico referido a impactos ("Generar las bases para que a la finalización del período de la administración Solís Rivera, pueda llevarse a cabo una evaluación de impacto del Programa"). Se reconoce que ese impacto no puede ser medible con la ejecución de esta evaluación, pero en ella se incluye una orientación y aproximación que posibilitará en el largo plazo determinar el impacto del Programa.
	¿Se diferencia entre los diferentes niveles de impacto?	Ninguna / No aplica No se diferencia entre niveles de impacto **Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)** Se diferencia con claridad los diferentes niveles de impacto	0 1 **5** 10	Aunque la evaluación no se centre en el análisis de los impactos, al abordar los hallazgos de la dimensión de diseño, en concreto la adecuación de la cadena de resultados, se indica lo siguiente: - *En relación con los efectos*: "se mantiene como principal efecto del Programa la mejora en la empleabilidad de los participantes, si bien, pueden establecerse categorías según el nivel educativo de los jóvenes y de otras de sus condiciones". (p. 57). "La inserción laboral que pueda producirse entre los participantes de Empléate una vez que finalizan su capacitación debe ser considerada, más bien, como un impacto del Programa y no tanto como uno de sus efectos". (p. 57) - *En relación con el impacto*: "el primero de los impactos que se considera es, precisamente, la inserción laboral de los jóvenes". "El resto de elementos relacionados con la inserción laboral (mantenimiento del empleo; empleo de calidad) adquieren la forma de una serie de condiciones encadenadas (figura 4)". (p. 57).
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica No se describe el contexto de la evaluación **Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención** Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	0 1 **5** 10	No hay un apartado específico dedicado a aportar claves contextuales que pueden haber influido en la realización de la evaluación. En todo caso, el epígrafe de "Limitaciones en el proceso de evaluación" incorpora algunas cuestiones que guardan relación con ello, fundamentalmente cuando se exponen las limitaciones derivadas de la metodología Aprender Haciendo, o cuando se señala la ausencia del responsable del programa en el proceso de evaluación. En cualquier caso, hay que advertir que la mayor parte de las limitaciones incluidas en ese epígrafe se centran en la disponibilidad de fuentes de información y de datos, más que al contexto en sí mismo.

3.2	Diseño y métodos	Calificación (texto)	(numérico)	Observaciones cualitativas
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (Los TdR no explicitan los intereses de la evaluación)	0	No se dispone de los TdR para tomar como referente cierto cuáles fueron los intereses (explícitos) del proceso evaluativo. En su defecto, habría que tomar como referente los objetivos de la evaluación, aunque estos se expresan de manera sintética y sin entrar en un gran nivel de detalle, señalándose como objetivo general "apoyar la toma de decisiones para la mejora del Programa", y como objetivos específicos las tres áreas que cubre esta evaluación: diseño, procesos y resultados (y un cuarto aspecto, los impactos, únicamente a modo de breve aproximación metodológica). Dado que el diseño evaluativo contempla además de las preguntas evaluativas las dimensiones del programa que se establecen en los objetivos, bajo un enfoque sistémico (contexto-estructura-diseño-procesos-resultados), puede decirse que el diseño se ajusta a esos intereses del proceso de evaluación. No obstante, dada la generalidad de los objetivos, no puede llegar a especificarse de manera muy concreta cuáles eran verdaderamente los intereses de la evaluación.
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	**5**	
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	10	
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica	0	Esta evaluación incluye diversa información de carácter metodológico que permite conocer qué se quiere analizar y cómo se va a obtener esa información, pero no ahonda en una reflexión de mayor calado sobre el diseño, los métodos de evaluación utilizados, o el porqué de las decisiones metodológicas adoptadas. En el cuerpo del Informe Final, existe un capítulo sobre "Caracterización y Metodología de la Evaluación", si bien al hablar de la metodología únicamente se señala que se recurrió a la metodología "Aprender Haciendo", y se incorpora una breve descripción de la misma, y posteriormente en el apartado de limitaciones se hace alguna valoración sobre la puesta en práctica de esa metodología. Ni el informe ni los Anexos llegan a explicitar ningún modelo metodológico de evaluación propiamente dicho, y no se desarrolla una fundamentación mínimamente sólida sobre por qué se eligió cada una de las técnicas, o cómo se complementan entre ellas. Una vez planteados los objetivos y el alcance de la evaluación, se definen directamente las preguntas de evaluación y los criterios con los que se vincula cada uno, sin ese respaldo de un modelo teórico que articule las diversas decisiones metodológicas adoptadas. Posteriormente se señalan las técnicas utilizadas, diferenciando entre fuentes primarias y secundarias. En uno de los Anexos se incorpora la matriz de evaluación, que incluye, además de esas preguntas y criterios, los aspectos a analizar, indicadores y fuentes de información. En los anexos también aparece un apartado dedicado a "Aspectos metodológicos", donde se hace explícita la finalidad de las técnicas utilizadas, pero de modo genérico y sin abordar la relación, triangulación o complementariedad entre ellas: <u>Entrevistas</u>: "(obtener) conocimientos e insumos que definan con claridad la situación del Programa desde su diseño, pasando por los procesos y los resultados obtenidos". <u>Encuestas</u>: "conocer de forma detallada datos, condiciones y aspectos del Programa". <u>Grupos focales</u>: "recoger las opiniones y las expectativas de los principales actores involucrados en el programa". <u>Observación</u>: "visibilizar aspectos relacionados con el trato, tiempos de atención, expectativas de los jóvenes y el valor de los medios utilizados para captar y seleccionar a la población del Programa". <u>Revisión documental y bases de datos</u>: "estudiar, conocer, describir y analizar la información contenida en documentos que dieron origen al programa, así como diferentes estudios y trabajos en torno al Programa".
		No se explicitan ni justifican el diseño y los métodos aplicados	**1**	
		Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados	5	
		Se explicitan y justifican suficientemente el diseño y los métodos aplicados	10	
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	Ninguna / No aplica	0	Esta información aparece en el Anexo 4, "Propuesta de Evaluación de Impacto para el Programa Empléate", con una extensión de 2 páginas, que pretende sentar las bases para una futura evaluación de impacto del Programa. En ese Anexo se apuesta por una evaluación de impacto de carácter cuasi-experimental, y se aportan algunas claves sobre cómo debería ser el grupo de comparación, así como las posibles estrategias para conformar ese grupo. También se indican los trabajos preparatorios que deberían ponerse ya en marcha para que esa evaluación de impacto sea factible, y se reconoce que para llevar a cabo una evaluación de este tipo con un cierto rigor sería necesario una inversión de recursos por parte del gobierno de Costa Rica.
		No se consideró lo contrafactual	1	
		Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró	**5**	
		Se consideró lo contrafactual y se especificó cómo	10	

				Si bien esa consideración de lo contrafactual resultaba bastante concreta y precisa, al estar concebida para su desarrollo futuro, y no aplicándose a la presente evaluación, se ha optado por valorar este aspecto con una puntuación intermedia.
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica Los métodos aplicados no cubren aspectos cuanti ni cuali Los métodos aplicados cubren aspectos cuanti o cuali **Los métodos aplicados cubren aspectos cuanti y cuali**	0 1 5 **10**	La evaluación se desarrolló con un enfoque mixto (cualitativo y cuantitativo). Los insumos cualitativos proceden de las entrevistas, grupos focales y observación, y los insumos cuantitativos de los tres tipos de encuestas realizados. A eso hay que añadir el marco contextualizador general que aporta la revisión bibliográfica, documental y de expedientes.
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) Entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica Se aplicó sólo un método Se aplicaron dos métodos **Se aplicaron una pluralidad de métodos (más de dos)**	0 1 5 **10**	Esta evaluación ha desarrollado diversas técnicas, tanto cuantitativas como cualitativas, habiendo tenido un considerable alcance sobre los diversos perfiles de informantes clave, como se puede ver en la siguiente tabla: Técnica / Método | Tipo | Nº Ediciones | Participantes Entrevistas | Semiestructuradas | 18 | 18 personas entrevistadas Grupos | Focales | 8 | 3 grupos, 33 gestores de empleo 3 grupos, 30 centros de formación 2 grupos, 14 beneficiarios Encuestas | Telefónicas | 3 | 950 beneficiarios 33 gestores de empleo 30 centros de formación Total = 1.013 participantes Observaciones | Directas, presenciales | 2 | Ventanilla de Dirección Reto Empléate GAM Expedientes | Revisión manual | 1 | Información sobre 2.822 beneficiarios reflejada en los 3.000 expedientes analizados (de los 8.000 existentes) Revisión bibliográfica y documental
3.3	**Procedimiento de selección**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección **Se explicitan en algunos casos (técnicas)** Se explicitan los procedimientos y criterios de selección	0 1 **5** 10	Los procedimientos y criterios de selección tuvieron que adaptarse a la realidad del Programa y a las limitaciones existentes. Por ejemplo, una de las limitaciones señaladas fue que las características de la población participante en el Programa Empléate (ej. carencias económicas para trasladarse, cambios de datos de contacto, falta de tiempo), suponían una dificultad para realizar los grupos focales con beneficiarios, por tanto hubo que acabar acudiendo directamente a los centros de formación para su celebración, y sólo se pudieron realizar dos. En el caso de las encuestas, se señala que se realizaron por vía telefónica por dos razones: falta de tiempo y recursos económicos. La encuesta a beneficiarios se realizó a una muestra, sin embargo en gestores de empleo y centros de formación se encuestó a la totalidad (33 y 30 respectivamente), indicándose que así se garantizaba la representatividad. En las entrevistas no se explicitan los criterios de selección de las personas concretas entrevistadas, más allá de una definición genérica de los perfiles (ej. funcionarios del Ministerio de Trabajo).
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras?	Ninguna / No aplica La muestra no es adecuada La muestra no es adecuada o el error no es aceptable **La muestra sí es adecuada y el error aceptable**	0 1 5 **10**	Los anexos recogen información clara y detallada sobre estas cuestiones, referidas a las 3 encuestas diferentes realizadas (beneficiarios, gestores de empleo, centros de formación). La encuesta a beneficiarios se aplicó a una muestra estadísticamente representativa de 950 participantes en el Programa. En los anexos se detalla, a modo de ficha técnica, diversos aspectos, como: población; marco muestral y método de muestreo (aleatorio estratificado con selección sistemática); muestra (prevista 948 y real 950); forma de administración; periodo del trabajo de campo; nivel de confianza (95%) y margen de error (3%). A la vista de toda esa información, puede señalarse que la muestra tiene un tamaño adecuado y un margen de error aceptable. Las otras dos encuestas (gestores de empleo y centros de formación) pudieron realizarse a la totalidad de la

				población (33 gestores y 30 centros, respectivamente), para garantizar la representatividad, lo cual es un logro muy satisfactorio.
3.4	**Análisis de datos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali **Se triangulan datos cuanti-cuali pero no de manera generalizada** Se triangulan datos cuanti-cuali de manera generalizada	0 1 **5** 10	A la hora de describir los hallazgos de la evaluación, se van exponiendo los resultados obtenidos a través de las diversas técnicas, según el aspecto del que se trate. En sentido estricto, más que una triangulación, se trata de una descripción de hallazgos o evidencias sobre una determinada cuestión, procedentes de las diversas fuentes, tanto cualitativas como cualitativas, pero sin que haya una intención expresa de realizar esa triangulación como elemento que refuerce el rigor y la robustez del ejercicio evaluativo. A modo de ejemplo, en el capítulo de hallazgos se encuentran frases como estas: "...según información obtenida tanto en los grupos focales como en las encuestas a los beneficiarios, se considera que..." "Existen divergencias en cuanto al proceso de capacitación, ya que tanto entre beneficiarios como empresarios hay opiniones encontradas".
3.4.2	¿Se cuantifican las respuestas, siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	Ninguna / No aplica No se cuantifican las respuestas cuando es posible **Se cuantifican las respuestas a veces cuando es posible** Se cuantifican las respuestas de manera generalizada cuando es posible	0 1 **5** 10	En el capítulo de hallazgos, generalmente se cuantifican las respuestas referidas a la técnica de la encuesta, pero no así las respuestas referidas al resto de técnicas, como las entrevistas o los grupos focales. Por tanto, esa cuantificación se centra en las técnicas cuantitativas, y no afecta a las técnicas cualitativas. Además, en los anexos se recoge de forma pormenorizada los resultados obtenidos en cada uno de los ítems de las tres encuestas realizadas (gestores de empleo; centros de formación; beneficiarios), lo que permite al lector poder conocer de forma exacta la magnitud alcanzada en cualquiera de esos ítems.
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica No se usaron datos del monitoreo Se usaron datos del monitoreo de manera insuficiente **Se usaron datos del monitoreo de manera suficiente**	0 1 5 **10**	Se ha concedido la puntuación más favorable a este aspecto, ya que aunque la información que existe sobre seguimiento es escasa, se ha utilizado y se ha incluido en la evaluación. Además, a lo largo del informe se percibe la importancia concedida a los elementos de monitoreo y seguimiento, y cómo esa necesidad llega a trasladarse a las recomendaciones. Entrando en más detalle, una de las limitaciones señaladas en el proceso de evaluación fue precisamente la insuficiente información de seguimiento: "El Programa no cuenta con un sistema propio de seguimiento debidamente estructurado, (…), con la excepción de la información que opera el Programa Nacional de Empleo (PRONAE), que cuenta con datos en una hoja Excel sobre los beneficiarios". En este sentido, el informe incluye una tabla de beneficiarios del Programa por sexo, según cantón (2012-2015), (Cuadro 3), siendo esta la única información procedente de la Base de Datos PRONAE, ya que los datos de jóvenes beneficiarios según especialidad no proceden de esa información de seguimiento, sino que se recabaron a través de la encuesta. También se indica como limitación que "los expedientes de los proyectos de capacitación son una importante fuente de información (…) pero esos datos no se encuentran digitalizados, ni sistematizados". Ante estas carencias, se revisaron de forma manual la información de 2.822 beneficiarios (procedentes de los expedientes), lo cual sirvió para calcular la muestra de la encuesta y para aplicar esos cuestionarios. Finalmente, hay que decir que una de las conclusiones de esta evaluación es que "el monitoreo y seguimiento se centra únicamente en aspectos de gestión general y de ejecución física-financiera", sin que haya seguimiento de resultados ni se explote la información contenida en los expedientes. Así, se llega a calificar a Empléate como "un programa ciego".
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica No se recurrió a fuentes externas Se recurrió a fuentes externas de manera insuficiente **Se recurrió a fuentes externas de manera suficiente**	0 1 5 **10**	El apartado del informe sobre el "Objeto de Evaluación" da comienzo con un epígrafe sobre "Descripción de la Problemática", que señala algunos datos de referencia sobre la formación y el desempleo juvenil a nivel mundial (OIT) y de América Latina y el Caribe (CEPAL). A continuación, se detallan diversos datos estadísticos concretos referidos a Costa Rica, procedentes de la Encuesta Nacional de Hogares (ENAHO) de 2010, y que ayudan a perfilar cuantitativamente la población objetivo del Programa Empléate.
3.5	**Criterios de evaluación**	**Calificación**		**Observaciones cualitativas**

3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	Dentro de la dimensión de "Diseño", una de las preguntas de evaluación guarda relación directa con la pertinencia: "¿Se ajusta el enfoque de Empléate a la problemática y a las necesidades de la población objetivo?" También se contempla responder a "¿En qué medida Empléate es complementario con otras iniciativas?", que se asocia a pertinencia y a complementariedad Por tanto, los temas de pertinencia están presentes en esta evaluación, y lo que se concluye es que el enfoque integral del Programa ofrece sólo un ajuste parcial a la problemática detectada y a las necesidades de la población objetivo. Por otro lado, cabe señalar que esta evaluación puso de manifiesto que la definición de la población meta contenía ambigüedades (ej. situación educativa), lo cual afectaba a la gestión del Programa y especialmente al alcance y naturaleza de los resultados esperables.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	5	
		Lo expone de forma concluyente y comprensible	**10**	
	b) logró las metas definidas	Ninguna / No aplica	0	La evaluación incluye, dentro de la dimensión de Resultados, una pregunta de evaluación que guarda relación directa con esta cuestión: "¿Consigue Empléate el propósito que persigue: Mejora de la empleabilidad?". A esta pregunta se le da respuesta comprensible y concluyente en el apartado de hallazgos, al hablar de la mejora en la empleabilidad de los participantes, en base a los datos arrojados por la encuesta a beneficiarios así como algún apunte procedente de las fuentes cualitativas. El nivel de detalle en la respuesta es considerable, ya que se incluye información referida a: - Jóvenes que finalizaron la capacitación y cuentan con empleo. - Jóvenes con empleo en áreas afines a la capacitación recibida. - Empresas aliadas que han empleado jóvenes del Programa. La conclusión a la que se llega a partir de tales hallazgos es que "El Programa Empléate cumple en forma parcial el propósito de mejorar la empleabilidad de los jóvenes", y diversas recomendaciones guardan relación con un mayor ajuste y eficacia del Programa.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	c) es eficiente	Ninguna / No aplica	0	Aunque la eficiencia se considera uno de los criterios y como tal está presente en algunas de las preguntas de evaluación, a la hora de exponer los hallazgos de la evaluación y las conclusiones/recomendaciones, no hay un bloque específico que aglutine y haga explícito todo aquello que está referido a la eficiencia del Programa. Lo que aparece, en su lugar, son alusiones relacionadas en mayor o menor grado con la eficiencia, y que salpican el texto, como: - "El tiempo promedio que se tarda en resolver una solicitud de ingreso al Programa, desde el momento en que el joven es preseleccionado puede ser utilizado para medir la eficiencia". - "Se recomienda que el proceso de intermediación e inserción tome en cuenta la información generada en los estudios de prospección anteriormente señalados, de forma que disponga de información suficiente para realizar de manera más eficiente y ágil la inserción laboral; esto se puede lograr asignando a las ventanillas de Empléate la función de intermediación". - "La prestación del servicio no es uniforme en todas las ventanillas donde se atiende población Empléate". - Algunas recomendaciones, como la 8 o la 10, influyen en la eficiencia. También hay un bloque de recomendaciones referidas a "ajustar y optimizar los recursos de Empléate".
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	
		Lo expone de forma concluyente y comprensible	10	
	d) causó impactos (outcome e impacto) deseados y no deseados.	Ninguna / No aplica	0	Una de las preguntas de evaluación, dentro de la dimensión de Resultados, es "¿Cuáles han sido los efectos no previstos de Empléate?". En sentido estricto, no se trata de impactos, sino de resultados/efectos, puesto que esta evaluación admite su dificultad o imposibilidad para llegar a medir impactos, dada la información disponible y el momento en que se encontraba el Programa. En cualquier caso, se indica que el Programa tiene efectos positivos, p.ej. la motivación para seguir estudiando posteriormente fuera del Programa; como efectos negativos, se apunta a que a veces esos egresados, por diversos factores, no consiguen un empleo, lo que les conduce a frustración y decepción. El informe incluye una figura que refleja gráficamente el nivel de motivación de los beneficiarios a través de las diferentes fases (Figura 8. Motivación del beneficiario según avanza el Programa). En definitiva, esta evaluación contempla el tema de los efectos no previstos, que se identifican fundamentalmente con la motivación, ya sea en sentido positivo o negativo. Pero no es uno de los ejes centrales de esta evaluación, por tanto
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	
		Lo expone de forma concluyente y comprensible	10	

				no puede decirse que la información sea concluyente y sirva para "dibujar" suficientemente el mapa de los impactos deseados y no deseados.
	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	Esta evaluación establece un objetivo específico referido a impactos ("Generar las bases para que a la finalización del período de la administración Solís Rivera, pueda llevarse a cabo una evaluación de impacto del Programa"). Se reconoce que ese impacto no puede ser medible con la ejecución de esta evaluación, pero en ella se incluye una orientación y aproximación que posibilitará en el largo plazo determinar el impacto del Programa.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	El Monitoreo y Seguimiento fue uno de los procesos analizados en esta evaluación, y como tal está claramente presente en los hallazgos, en las conclusiones y en las recomendaciones. La evaluación concluye que "este proceso presenta serias debilidades, dado que el tipo de seguimiento que se realiza se limita a monitorear un indicador basado en la cantidad de jóvenes atendidos por el Programa y los recursos ejecutados"(gestión general y ejecución física-financiera). Por tanto, llega a considerarse que Empléate "se limita a acumular expedientes sin automatizar la información", lo cual le convierte en "un programa ciego". Además, tampoco es posible hacer análisis comparativos, ya que no es posible distinguir entre las tres modalidades del programa. Tampoco existe ningún seguimiento referido a la inserción del joven al mercado laboral, ya que ni siquiera esta área de intervención ha sido abordada suficientemente por el Programa. A nivel de recomendaciones, se indica claramente la necesidad de "establecer un plan de monitoreo y seguimiento del Programa Empléate, que facilite un mayor conocimiento del programa y sus resultados y puedan utilizarse para apoyar la toma de decisiones", así como "establecer un Sistema de información (en Excel o similar) para incluir los datos referentes al programa que facilite la utilización de la información y facilite la toma de decisiones".
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?		0	Aunque ya se ha señalado que esta evaluación no cubre en sentido estricto un análisis de impactos, el apartado del Informe que en mayor medida podría aproximarse a ese aspecto de los "cambios observados" serían los hallazgos de evaluación referidos a la dimensión de resultados, en concreto a la mejora de la empleabilidad de los jóvenes, de forma general, y de modo más pormenorizado en tres perfiles: - Jóvenes que finalizaron la capacitación y cuentan con empleo. - Jóvenes con empleo en áreas afines a la capacitación recibida. - Empresas aliadas que han empleado jóvenes del Programa. En algunas ocasiones, esa descripción de efectos o cambios se pone en relación con elementos del Programa, pero no es lo habitual, ni es algo que se refleje de forma sistemática. Ejemplo: - "El 27% de los jóvenes que se encuentran trabajando, lo hacen en labores que tienen mucha relación con la capacitación recibida; lo anterior indica que un porcentaje considerable de jóvenes están siendo capacitados en especialidades que no son las demandadas por el mercado. Situación que podría verse solventada realizando estudios de prospección, mediante los cuales sería posible conocer las demandas del mercado y a los beneficiarios se les facilitaría conseguir empleo una vez finalizada la capacitación". (p. 98). Más bien, los cambios o efectos se ponen en relación con elementos del entorno, que trascienden al Programa: - "…el 50% no logran obtener un empleo, independientemente del área en que fueron capacitados, a pesar de optar por varios empleos. Lo atribuyen a dos razones principales: falta de experiencia laboral y escasa oferta de empleo en la zona". (p. 99).
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	**5**	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	

4.	Calidad del informe y utilidad	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	Los hallazgos de evaluación dan respuesta de forma adecuada a las trece preguntas de evaluación planteadas, y las conclusiones y recomendaciones también guardan coherencia con esa estructura. Algunas preguntas, como la de "¿Existen funciones o procesos que el Programa no está realizando?", tienen un abordaje más sintético, si bien esto no constituye en sí mismo una debilidad o problema, sino que resulta lógico que la respuesta a algunas preguntas, por su contenido, no requiera un desarrollo tan amplio como otras.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	El informe señala explícitamente que la evaluación "busca identificar fortalezas y debilidades de la intervención, con el fin de proponer medidas correctivas para mejorar y facilitar la implementación del Programa". En líneas generales, el informe pone más el acento en las debilidades o elementos a mejorar (quizás de modo justificado, por requerir el Programa considerables mejoras), que en los puntos fuertes o aciertos del Programa. Sobre aspectos positivos, se alude por ejemplo al efecto motivador del Programa entre los jóvenes (aunque en realidad esa motivación se despliega en diversas fases, y puede conducir posteriormente a una frustración por incumplimiento de las expectativas); o a las fortalezas adicionales que poseen los gestores de empleo y que facilitan la prestación del servicio (ej. experiencia). Por otro lado, esa visión crítica también está presente cuando se detallan las diversas limitaciones que ha tenido el proceso de evaluación, y el lugar que ocupan éstas dentro del Informe (justo después de la metodología y antes de exponer las preguntas de evaluación y las técnicas utilizadas). En definitiva, el Informe se centra especialmente en identificar, fundamentar y explicar las debilidades del Programa, lo cual resulta valioso para la evaluación, si bien se echa en falta, de forma paralela y equilibrada, una mayor identificación de las fortalezas o de los elementos que se han mostrado más positivos en el mismo.
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	**5**	
		Se describen balanceadamente puntos fuertes y débiles	10	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	En líneas generales, el Informe integra en su contenido tanto el plano descriptivo como el plano valorativo, y lo hace de forma clara y vertebrada, aunque no haya una mención expresa bajo esa nomenclatura de "descriptivo "y "valorativo". Existe una diferenciación clara, a nivel de capítulo, entre los hallazgos, las conclusiones y las recomendaciones, de modo que en los hallazgos se exponen los resultados obtenidos a partir de las distintas fuentes y técnicas, señalando habitualmente en cuál de esas fuentes o técnicas se basa cada una de las afirmaciones. Por su parte, en las conclusiones se va avanzando desde lo descriptivo hacia elementos de corte más valorativo/conclusivo, mientras que finalmente las recomendaciones aportan orientaciones sobre la dirección a tomar o la acción a emprender ante las necesidades o déficits que se han ido manifestando a lo largo del proceso de evaluación.
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	**10**	
4.4	¿Se especifica cómo las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	Aunque no logra ser un aspecto que se contemple de manera generalizada o sistemática en esta evaluación, cabe decir que el apartado del Informe donde se refleja en mayor medida esta atención a los aspectos de calidad, fiabilidad y validez de la información, corresponde a "Limitaciones en el proceso de evaluación". Ahí, se realizan matizaciones como las siguientes: - *Manejo prudente de Información de centros de formación*: "…el equipo evaluador no logró verificar la veracidad de la información suministrada: considerando que son actores beneficiados por el programa, su opinión puede ser subjetiva. Por esta razón, el equipo evaluador manejó prudentemente esta información". En otro momento del Informe, se subraya también que "es necesario y urgente diseñar un plan de monitoreo del programa que le posibilite disponer de información confiable para la toma de decisiones oportunas y facilitar la mejora continua del Programa". En definitiva, se plasma una cierta preocupación porque los datos sean confiables, pero existe todavía un notable margen de mejora al respecto en el proceso de evaluación.
		No se especifica	1	
		Se especifica en ocasiones	**5**	
		Se especifica en general	10	

4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica	0	Se concede la puntuación máxima ya que esta evaluación consigue presentar un capítulo de conclusiones sintético, claro, y que guarda relación directa con los resultados expuestos en el capítulo de hallazgos de evaluación, que a su vez se vinculaban a las preguntas de evaluación.
		Las conclusiones no están basadas en los resultados	1	
		Alguna o algunas conclusiones no están basadas en resultados	5	
		Todas las conclusiones están basadas en los resultados	**10**	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	El informe de evaluación manifiesta una coherencia y solidez suficiente en su estructura, que hace que las recomendaciones puedan ser leídas adecuadamente como fruto de las conclusiones, con la sistematización propia que ha de conllevar un proceso de evaluación.
		No se llega ni de forma lógica ni comprensible	1	
		Se llega de forma lógica o comprensible pero no las dos	5	
		Se llega de forma lógica y comprensible	**10**	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	Se establecen 20 recomendaciones, y todas ellas llevan asociadas una institución o entidad responsable de su implementación. Esas recomendaciones se clasifican según guarden relación con el diseño, los procesos o los resultados, por tanto son relevantes ya que esa era la estructura sobre la que pilotaba desde el principio esta evaluación. Asimismo, las recomendaciones resultan suficientemente específicas, a lo que contribuye que cada recomendación vaya acompañada de algunos párrafos explicativos que concretan el enfoque o el tipo de acciones a realizar.
		No son relevantes ni específicas	1	
		Son relevantes o específicas pero no las dos	5	
		Son relevantes y específicas	**10**	
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	El resumen incluido en el Informe Final está adecuadamente estructurado y es comprensible. Tiene una extensión aproximada de 9 páginas, y se centra sobre todo en exponer las conclusiones y recomendaciones, especificándose en estas últimas quién tiene la responsabilidad de su traslación a la realidad del Programa. Aun siendo esa su parte más destacada, el Resumen ejecutivo realiza también una breve introducción inicial sobre el Programa, así como: una explicación general de la evaluación; las preguntas de evaluación; la exposición de las diferentes técnicas utilizadas, y su magnitud (ej. nº de entrevistas).
		El Informe no contiene un resumen	1	
		El informe contiene un resumen estructurado	5	
		El informe contiene un resumen estructurado, bien legible/comprensible	**10**	
5.	**Exhaustividad y adecuación de la documentación**			
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica	0	Los TdR no están disponibles en los anexos, tampoco forman parte de la documentación consultada en esta metaevaluación, ni se encuentran disponibles ni accesibles en la web.
		Los TdR no son accesibles en anexos	**1**	
		Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	
		Los TdR son accesibles en la lengua común de trabajo o en una traducción	10	
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	En los Anexos se realiza una descripción de aspectos metodológicos (Anexo 21), describiendo cada una de las técnicas utilizadas. Pero no aparece ningún tipo de cronograma de la fase de campo, ni se indican de forma expresa qué personas fueron entrevistadas o participaron en los grupos focales. Sí se indica, sin embargo: - Los centros de formación que participaron en la encuesta. - Las instituciones a las que pertenecen los gestores de empleo que participaron en la encuesta. - El Reto al que se acudió, dentro de la técnica de Observación. (Reto GAM).
		No se incluyen en anexos	1	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	**5**	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	El informe final se cierra con el capítulo de Bibliografía, y los Anexos recogen de forma detallada las diferentes herramientas utilizadas: - Cuestionario para Gestores de Empleo, y resultados obtenidos. - Cuestionario para Centros de Formación, y resultados obtenidos. - Cuestionario para beneficiarios, y resultados obtenidos. - Diversas guías de entrevista (personeros de la AED; DESAF; superiores de gestores de empleo; empresas que contratan a beneficiarios de Empléate; experto en Sistemas de Información; Instituciones IMAS, CPJ, INAMU, INA; Unidad Ejecutora de Empléate; especialista en temática laboral. - Grupos focales con beneficiaros de Empléate.
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	**10**	

				- Grupos focales con centros de formación. - Grupos focales con gestores de empleo.
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos?	Ninguna / No aplica	0	El informe está estructurado de manera adecuada y su lectura se facilita con tablas y gráficos, aunque no presenta resúmenes de los capítulos.
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	5	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	**10**	

4.4.8 Evaluación 08: OAPVD

0.1 Nombre del proyecto / programa evaluado	**PROGRAMA DE ATENCIÓN DE LA OFICINA DE ATENCIÓN Y PROTECCIÓN A LA VÍCTIMA DEL DELITO (OAPVD).**
0.2 Institución que implementa	Ministerio Público-Poder Judicial.
0.3 Evaluadores	Se trata de una Investigación evaluativa en el marco de la cooperación académica de la Maestría en Evaluación de Programas y Proyectos de Desarrollo (MEPPD) de la Universidad de Costa Rica y el Centro para Evaluación (CEval) de la Universidad del Sarre, Alemania. Según consta al comienzo del Informe, finalmente la composición del equipo fue la siguiente: Equipo de evaluación: Andrea Meneses Rojas (coordinación); Carolina Zúñiga Zamora; Susanne von Jan; Carmen Marín González; Fainier Candamo Barrantes. Con la colaboración de: - CEval: Dr. Prof. Reinhard Stockmann (Director) - Integrantes de la Comisión de la MEPPD-UCR: M.Sc. Olman Villarreal (Director MEPPD), M.Sc. Nidia Morera (Directora Escuela y Posgrado de Trabajo Social), Dra. Mayela Cubillo (Directora Centro de Investigación y Capacitación en Administración Pública-CICAP). - Docentes de la Universidad de Costa Rica: M.Sc. Corinna Schopphoff, M.Sc. Ana Victoria Molina, M.Sc. Johanna Alarcón Rivera. Con el apoyo de: - Pasantes de la Maestría en Evaluación, Universidad del Sarre: Tania Jensen, Katrin Windolf, Sandra Schopper. - Investigadores/as de la Universidad del Sarre: Jessica Gödert, Simeón Schächtele, Dietmar Mälzer, Sandra Baethge. - Asistentes de la MEPPD-Universidad de Costa Rica: Luis Fernando Castillo Vindas, Xinia Mora Valverde. - Asistentes de CEval-Universidad del Sarre: Lennart Schemeling, Kristin Alle. - Escuela de Trabajo Social – UCR. - Centro de Investigación y Capacitación en Administración Pública-UCR. - Escuela de Administración Pública-UCR. - Programa Fomento de Capacidades en Evaluación (FOCEVAL): GIZ y DEval. - Sistema de Estudios de Posgrado - UCR: Dra. Cecilia Díaz Oreiro (Decana).
0.4 Monto del proyecto/ programa	Dentro de los resultados de evaluación, en el bloque sobre recursos financieros, se incluyen datos cuantitativos sobre el presupuesto anual destinado al conjunto de la Oficina, pero no específicamente al Programa de Atención, que es lo que se evalúa en concreto. (Gráfico nº7. Historial del presupuesto OAPVD después de la aprobación de la Ley 8720, según año (millones de colones), p. 158).
0.5 Monto para la evaluación	No se refleja esta información.
0.6 Descripción del proyecto/programa evaluado	
La evaluación realizada se denomina "Evaluación de la calidad del servicio del Programa de Atención de la Oficina de Atención y Protección a la Víctima del Delito (OAPVD), del Ministerio Público, durante los años 2010, 2011, 2012, 2013" . La Oficina de Atención y Protección a la Víctima del Delito (OAPVD) se crea en el año 2000. Actualmente cuenta con más de 100 funcionarios/as y se encuentran ubicados en las diferentes provincias del país. Cuenta con dos programas principales, el Programa de Atención y el Programa de Protección. Esta evaluación se ha centrado en el Programa de Atención, que estructura esos procesos de atención en tres fases (inicial, seguimiento y cierre), y en tres áreas de atención (trabajo social, psicología, legal).	
0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL	
Además de estar incluidos en la relación de personas o instituciones involucradas en la evaluación y en los agradecimientos, en diversas páginas del Informe se hace mención al apoyo brindado a través del programa FOCEVAL, por ejemplo ello permitió la contratación de la persona coordinadora. En el capítulo de lecciones aprendidas y buenas prácticas, se incluyen algunas consideraciones relacionadas con ese apoyo por parte de FOCEVAL, de cara a incorporar mejoras futuras, y relacionadas fundamentalmente con la disponibilidad de tiempo efectivo para participar en las tareas evaluativas: - "Se debe evitar grupos grandes de trabajo, donde existan más de 5 personas. La presente evaluación, por cuestiones presupuestarias, llegó a contar con un equipo de evaluación de 7 personas con tiempos limitados, además de contar con el apoyo en diferentes tareas y actividades de hasta 10 personas adicionales para permitir el avance de la evaluación. Lo anterior generó el recargo de funciones en algunas personas, e intensificando las labores de comunicación de la coordinadora".	

- "Si bien una fortaleza del equipo de evaluación fue la multidisciplinariedad y compromiso de las personas que lo integran, así como su experiencia probada en diferentes áreas de trabajo (trabajo social, evaluación, administración pública, sociología, temas judiciales), los tiempos de dedicación a la evaluación, por falta de recursos, fue muy limitada. Las futuras evaluaciones realizadas desde la MEPPD con la cooperación internacional, deben necesariamente contar con dos personas como mínimo que se dediquen tiempo completo a la evaluación del programa o proyecto".

Por otro lado, en lo referente a la socialización de resultados, se indica que el Informe se difundió en diferentes espacios, por ejemplo entre los integrantes del grupo de coordinación de FOCEVAL.

Finalmente, señalar que comparativamente el Informe se detiene con más detalle en la relación entre el equipo evaluador y la jefatura y personal de la OAPVD (objeto evaluado), que en la relación entre equipo evaluador y FOCEVAL.

0.8 Cualificación cuantitativa (con comentarios) según criterios			
71,6 [Sobre 100]	**7,5**	***Imparcialidad e independencia:*** [Sobre 10]	A pesar de que los propios responsables y personal de la OAPVD colaboran en la preparación y realización de la evaluación, se consigue en buena medida los objetivos de imparcialidad e independencia, ya que el equipo evaluador es quien asume las riendas y toma las decisiones metodológicas, además de utilizarse diversas técnicas y fuentes de información.
	8,8	***Equidad, participación y transparencia:*** [Sobre 10]	No se detectan problemas en plantear una evaluación comprensible y en dar a conocer sus resultados, así como numerosos detalles acerca de su implementación. Sin embargo, como debilidad aparece que apenas hay implicación de otros grupos diferentes a las propias OAPVD.
	39,3	***Credibilidad y exactitud:*** [Sobre 50]	Un considerable número de ítems englobados en este criterio reciben la máxima puntuación. Puede decirse que esta evaluación presenta calidad metodológica, y tiene en cuenta los diversos aspectos necesarios para aportar credibilidad y exactitud (descripción de objetivos y objeto, análisis de grupos, elaboración de un modelo de evaluación, desarrollo y explicación sobre las técnicas empleadas, etc.) Por la propia naturaleza y temática de esta evaluación, se analizan en mayor medida los efectos (efectos directos) que los impactos propiamente dichos.
	11,8	***Calidad del informe y utilidad:*** [Sobre 20]	La elaboración del informe resulta cuidada, se intenta ofrecer la mayor información posible, si bien eso mismo supone un elemento de mejora, ya que todo eso no viene acompañado de un esfuerzo de síntesis, no se destacan las ideas más importantes ni aparece un resumen ejecutivo. Por tanto, el lector tiene que realizar un esfuerzo adicional para localizar la información más relevante o de su interés. Es por ello que este criterio recibe la puntuación más baja.
	4,3	***Exhaustividad y adecuación de la documentación:*** [Sobre 10]	A lo largo del informe y de los anexos se intenta ofrecer la mejor y mayor información posible, no sólo de los resultados de la evaluación, sino de su propio desarrollo metodológico y analítico. Aun así, se echa en falta alguna documentación específica, como el detalle del cronograma o el listado completo de informantes que han participado en la evaluación.

0.9 Resumen de fortalezas de la evaluación.

- La evaluación contiene gran cantidad de información, que una vez procesada y sintetizada por quienes encargaron la evaluación, puede ser de gran interés y utilidad para aplicar al Programa.
- Siendo una evaluación académica, se presta especial atención a elementos conceptuales, aclaraciones terminológicas, alusiones bibliográficas dentro del texto.
- Se refleja de manera expresa el enfoque de evaluación utilizado, que equivale al modelo o teoría de evaluación, y posteriormente el marco evaluativo concreto (diferentes elementos de la matriz).
- Importantes esfuerzos en el desarrollo de análisis de efectos (ej. se elabora un índice de empoderamiento).
- Se vuelcan interesantes reflexiones en el capítulo de lecciones aprendidas y buenas prácticas. En general, a lo largo del Informe aparecen diversas claves sobre el "hacer evaluativo", que resulta de utilidad a los profesionales de la evaluación y a las personas interesadas en conocer con más detalle el proceso de implementación de las evaluaciones.

0.10 Resumen de debilidades de la evaluación.

- En bastantes ocasiones, la extensión en el desarrollo de los contenidos dificulta obtener una idea global; se echa en falta un mayor esfuerzo de síntesis y de resaltar aquello que el equipo evaluador considera más importante. En general, la profundidad y forma como está estructurada la información no facilita la labor de lectura.
- Ausencia de un capítulo específico sobre conclusiones. Éstas aparecen entremezcladas en el capítulo de resultados de la evaluación y en el de recomendaciones. Eso dificulta obtener una visión global y sintética sobre las principales ideas aportadas por la evaluación.
- Elevado número de recomendaciones (que aparecen, además, sin numerar ni priorizar, y se repiten dos veces, ya que aparecen tanto en el capítulo de resultados como en el de recomendaciones). Este hecho puede dificultar su posterior implementación. Además, tienen un nivel de detalle desigual, y en ocasiones no está claro de que se trate de una recomendación propiamente dicha, puesto que se entremezcla con elementos de reflexión en general.
- No hay una participación significativa de las personas usuarias en el diseño o implementación de la evaluación, simplemente adoptan la figura tradicional de "informantes clave".

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	Como se refleja en el capítulo 1 (contexto de la evaluación), se establece una relación de cooperación entre la Maestría de la Universidad de Costa Rica y la OAPVD, con la implicación también de otros actores, como CEval y FOCEVAL. Es decir, hay personal de la propia Oficina implicada en las labores de preparación y apoyo a la realización de la evaluación. Es por ello que se ha calificado como evaluación mixta si bien se aprecia que el liderazgo del proceso evaluativo lo asume la Universidad de Costa Rica a través de la Maestría, más que la propia OAPVD.
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	**5**	
		Evaluación externa (100% externo a instituciones evaluadas)	10	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados?	Ninguna / No aplica	0	Las técnicas empleadas para recoger información fueron diversas (talleres de trabajo; cuestionario autoadministrado; entrevistas, observación), lo que conllevó la captación de una pluralidad de perspectivas. Asimismo, en el análisis de datos, al realizarse de forma bastante explícita la triangulación, también queda cubierta esa pluralidad y complementariedad. Por último, y esto es un valor añadido de esta evaluación, en la presentación de resultados se procuró contar previamente con la opinión y matizaciones de los principales agentes implicados, a través de un taller previo de validación, y posteriormente un taller de difusión del informe: - *"Una lección aprendida con respecto a este tema, es utilidad que tienen los talleres de conclusión de la evaluación. El taller previo de validación de resultados y recomendaciones con el cliente de la evaluación, así como el taller final de difusión del informe le dan legitimidad al proceso evaluativo, propiciando una mejor aceptación de la información contenida. Parte de la acogida de esta evaluación fue gracias a la discusión previa que se realizó con funcionarios, funcionarias y tomadores de decisión del programa". (p. 212)*
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	5	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos)	**10**	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos?	Ninguna / No aplica / No hay información en los documentos consultados	0	En esta evaluación no hubo implicación de los grupos meta (personas usuarias), sino principalmente estuvieron involucrados la jefatura de la OAPVD y las personas funcionarias de las diversas oficinas implicadas. Es por ello que se le otorga una puntuación intermedia en este ítem, si bien se reconoce un esfuerzo importante en implicar a estos funcionarios (gestores/ejecutores del programa) en diferentes momentos del proceso. Como se señala al principio del Informe, al describir el contexto del proyecto evaluativo: - "Entre los meses de septiembre del 2012 y febrero del 2013, se realizan varias sesiones de trabajo entre el equipo evaluador MEPPD- UCR y los/as funcionarios/as de la OAPVD, en donde se toman acuerdos importantes y se desarrollan actividades claves para iniciar la definición de la evaluación. Dentro de esas actividades sobresalen: la reconstrucción de la Teoría del Programa de la OAPVD, la realización de un taller de trabajo para desarrollo de un Mapeo de Actores de la OAPVD y de la Evaluación; y por último una visita de representantes de CEval y del equipo de evaluación a la OAPVD, para aclarar aspectos básicos sobre el acceso a la información durante el proceso evaluativo (principalmente referido a libros electrónicos y expedientes de víctimas". (p. 16). También se señala que una vez finalizada la evaluación del Programa de Atención de la OAPVD, se discutieron los principales hallazgos y recomendaciones con funcionarios/as del Programa de Atención durante un taller de conclusión, antes de realizar una presentación final donde se incluyeron otros Stakeholders incluidos dentro del proceso evaluativo (no se detallan cuáles). Por otra parte, se compartió en diferentes espacios las principales implicaciones teóricas y metodológicas de la evaluación, así como las lecciones aprendidas del proceso.
		No hay participación	1	
		Participan escasa o desigualmente en planificación o implementación	**5**	
		Participan activamente en la planificación y en la implementación	10	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	El apartado 3.2 recoge de manera clara y sintética los Interrogantes de la Evaluación (cuadro nº2, p. 14), y lo hace en el lugar que corresponde, después de haber expuesto el contexto evaluativo, las características del objeto a evaluar, y los objetivos de la evaluación.
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	5	
		Las preguntas son claras y adaptadas	**10**	

2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica**	**0**	No se dispone de los TdR, si bien los interrogantes de evaluación, que son un elemento central en la articulación de la evaluación, resultan suficientemente claros y comprensibles.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	Como parte de las actividades programadas previo al trabajo de campo, se coordinó un taller de capacitación/sensibilización con funcionarios/as de la OAPVD, para que el personal conociera los objetivos y alcances de la evaluación, así como la relevancia de su participación en el proceso evaluativo. Finalmente, se contó con la participación de 26 funcionarios/as del Programa de Atención de diversas regiones, incluyendo las jefaturas. Entre otros aspectos, el taller resultó esencial para contar con la participación asertiva de las personas funcionarias en las etapas siguientes de la evaluación. Además, en este taller se mostró la preocupación de los/as funcionarios/as sobre la manera de promover la participación de los/as usuarios en las entrevistas, por lo que se discutieron algunos beneficios o facilidades que la OAPVD podía ofrecer para incentivar esa participación. (p. 55 Informe Final).
		No hubo un taller de inicio con stakeholders para facilitar su participación	1	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	
		Hubo un taller de inicio con stakeholders	**10**	
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica	0	Se entiende por accesibilidad su disposición a través de internet. En este sentido, no se ha localizado en Internet el informe ni los anexos de esta evaluación.
		No son accesibles para el público (ni Informe ni resumen)	**1**	
		Son accesibles para el público (Resumen o extracto)	5	
		Son accesibles para el público (Informe y resumen)	10	
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	El capítulo 3, sobre Diseño Evaluativo, se inicia con el epígrafe 3.1., que aborda los "Objetivos de la Evaluación". Al tratarse de un proyecto en el marco de una cooperación académica entre instituciones, se contemplan tanto objetivos académicos, como propiamente objetivos evaluativos. El cuadro nº 1 (p. 43 y 44), recoge de manera pormenorizada esos objetivos de evaluación, distinguiendo entre generales y específicos, y entre evaluación de procesos (formativa) y orientada a efectos (sumativa).
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	
		Se describe el objetivo de la evaluación de manera imprecisa	5	
		Se describe de manera precisa el objetivo	**10**	
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica	0	Se dedica un capítulo de aproximadamente 25 páginas a este aspecto (Capítulo 2. Programa a evaluar: la OAPVD del Ministerio Público). En él, se abordan distintas cuestiones: - Referentes jurídicos. - Estado costarricense y algunas tendencias del Sistema de Justicia. - Contexto institucional de la OAPVD (Poder Judicial). - Teoría del Programa de la OAPVD. - Tipos de finalización del proceso penal (complemento a la teoría del programa). - Ruta crítica de intervención formal del Programa de Atención ("tipo ideal" de atención). Asimismo, algunos de los Anexos ofrecen información sobre el objeto de la evaluación, como: mapeo de actores; marco conceptual; información adicional sobre procesos de atención; marco comparativo de oficinas de atención a nivel latinoamericano. En definitiva, el tratamiento del objeto de evaluación es extenso y pormenorizado (quizás incluso puede llegar a resultar un tanto abrumador para aquellos lectores poco familiarizados con este tipo de programas y servicios, por lo cual podría haber sido conveniente incluir una ficha resumen con la información más relevante al respecto). Por último, cabe señalar que la información financiera sobre el objeto evaluado aparece en el capítulo de resultados, y no tanto en el capítulo de descripción del objeto evaluado. Seguramente, ello se debe a que la información disponible es escasa, y se utilizó directamente para responder a uno de los interrogantes de la evaluación, referido a recursos financieros.
		No se describe el objeto de evaluación	1	
		Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes	5	
		Se describe con precisión (al menos los ítems contemplados)	**10**	
3.1.3		Ninguna / No aplica	0	

	¿Se realizó un análisis de los grupos de interés (stakeholder)?	No se realizó un análisis de stakeholders Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.) **Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)**	1 5 **10**	Dentro del Capítulo 3, sobre Diseño Evaluativo, el epígrafe 3.6 está referido a "Participantes y Destinatarios de la Evaluación", donde se indica: - *"La evaluación en sus diferentes momentos contó con la participación de distintos grupos de personas involucradas que se denominan actores clave de la evaluación, a su vez también se delimitan los destinarios de la evaluación, es decir, las personas interesadas en la información que genere la evaluación para su uso ya sea para la toma de decisiones (nivel interno) o para la generación de conocimiento (nivel externo)".* También hay que señalar que en el Anexo 3 se incluyen los Mapeos de Actores, por un lado el mapeo de actores OAPVD (se entiende que equivale a los actores clave de la evaluación), y por otro el mapeo de actores de la evaluación OAPD (se entiende que equivale a destinatarios de la evaluación). Estos mapeos se exponen de manera gráfica, distinguiéndose varios niveles en cada uno de ellos, según grado de importancia: veto, calves, primarios y secundarios.
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica No se elaboró Se elaboró pero no se usa en la evaluación o no es concluyente **Se elaboró, se usa en la evaluación y es concluyente**	0 1 5 **10**	En esta evaluación se reconstruye la teoría del programa de la OAPVD, así como la ruta crítica de intervención ("tipo ideal" de atención), con el fin de determinar cómo se está respondiendo desde el Programa a las necesidades y demandas de las personas víctimas del delito. Todo este tipo de información no se encontraba disponible en las fuentes documentales oficiales del Programa, tuvo que ser reconstruida en el marco de los trabajos realizados. Como particularidad de esta evaluación, la reconstrucción de la teoría del programa del conjunto de la Oficina permitió identificar qué programa tenía las mejores condiciones para ser evaluado. Como se indica en la p. 125: - "Se utilizó la técnica de Cadena de Resultados, la cual fue de mucha utilidad para poder relacionar causalmente cómo el Programa de Atención se propone empoderar a la víctima del delito para que tenga una participación más activa en el proceso penal y disminuir la revictimización". Cuando en el capítulo 5 se abordan los resultados de la evaluación, es decir, los hallazgos, se incluye un epígrafe dedicado a los efectos del Programa. Ese análisis de efectos se concentró en tres aspectos: 1) el empoderamiento de las víctimas, 2) la disminución de la revictimización y 3) la incidencia del programa en la forma de finalización del proceso penal. Por último, hay que señalar que algunos Anexos incluyen información adicional relacionada con la teoría de intervención, cadena de resultados o lógica de impacto: - Anexo 4. Teoría del Programa (donde se presenta un gráfico completo con insumos, actividades, productos, aprovechamiento de productos, efectos directos e impacto). - Anexo 9. Información adicional del capítulo efectos. En definitiva, puede decirse que en esta evaluación hay un abordaje profundo y suficientemente aplicado de los aspectos contemplados en este ítem, por ello obtiene una alta valoración.
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica No se formularon Se formularon pero no son concluyentes **Se formularon y son concluyentes**	0 1 5 **10**	En primer lugar, hay que recordar que esta evaluación se llevó a cabo bajo el enfoque del modelo CEval, conocido como Modelo de Impacto, así como los aportes del enfoque de la gerencia social. Esta evaluación ofrece un considerable detalle y nivel de desarrollo en lo que a análisis de efectos se refiere. En cuanto a impactos, la teoría del programa aporta cierta información, aunque de carácter general. El principal impacto que se define es: "Un sistema de justicia transparente, eficiente, accesible y de resultados, para contribuir a la solución de conflictos sociales y a la disminución de la impunidad" (Anexo de Teoría del Programa).

	¿Se diferencia entre los diferentes niveles de impacto?	Ninguna / No aplica No se diferencia entre niveles de impacto **Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)** Se diferencia con claridad los diferentes niveles de impacto	0 1 **5** 10	La diferenciación en niveles, o más bien en temáticas de incidencia, se realiza en los efectos directos, no puede decirse propiamente que haya diferenciación en niveles en impactos.
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica No se describe el contexto de la evaluación **Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención** Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	0 1 **5** 10	A lo largo del informe se incluyen elementos que contribuyen a una descripción suficientemente detallada del contexto de la evaluación, fundamentalmente en el capítulo 1, titulado "Contexto del proyecto evaluativo" (5 pág. aprox.), donde se van describiendo las vicisitudes y circunstancias que han rodeado a esta evaluación. De manera resumida, el contexto que originó esta demanda de evaluación fue así: - *"La necesidad de evaluar el servicio brindado proviene de la Oficina de Atención misma, quien observó que sus procesos de trabajo debían ser revisados y analizados para identificar las áreas de mejora a trabajar en el futuro. La Maestría en Evaluación de Programas y Proyectos de Desarrollo aprovecha la oportunidad de colaborar con una entidad pública, utilizando recursos de diferentes instituciones y entidades académicas, pero con el apoyo técnico del Centro para Evaluación de la Universidad del Sarre, en Alemania y la colaboración financiera del Servicio Alemán de Intercambio Académico. Con estas entidades se firmó un programa de cooperación desde el año 2008, cuya vigencia finaliza a finales del 2015"*. (Informe, p.11). Sin embargo, a pesar de estas explicaciones sumamente detalladas, no se observa que ese contexto se ponga en relación con la teoría de intervención del programa.
3.2	**Diseño y métodos**			
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (Los TdR no explicitan los intereses de la evaluación) El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe) El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe) **Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)**	**0** 1 5 **10**	Esta evaluación ofrece un abordaje metodológico bastante cuidado y detallado, tanto en el capítulo 3 (Diseño Evaluativo) como en el Anexo 6 (Marco Evaluativo), queda reflejada adecuadamente la correspondencia entre los interrogantes de evaluación (que equivaldría a los intereses evaluativos) y las dimensiones de análisis, variables, indicadores, técnicas y fuentes.
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica No se explicitan ni justifican el diseño y los métodos aplicados Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados **Se explicitan y justifican suficientemente el diseño y los métodos aplicados**	0 1 5 **10**	Como se ha señalado en el punto anterior, esta evaluación ofrece un abordaje metodológico bastante cuidado y detallado, tanto en el capítulo 3 (Diseño Evaluativo) como en diversos anexos. Además, una de las fortalezas de esta evaluación es que no sólo vincula los interrogantes de evaluación con los indicadores y técnicas/fuentes, sino que describe el enfoque o modelo de evaluación que sirve de marco global conceptual-interpretativo a los trabajos realizados en la evaluación.
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	Ninguna / No aplica No se consideró lo contrafactual Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró **Se consideró lo contrafactual y se especificó cómo**	0 1 **5** **10**	En el apartado 3.7.1 del Informe final, sobre Diseño Evaluativo, se explica detalladamente cómo se abordó este asunto, y las decisiones que hubo que adoptar (p. 51 y ss.): - El diseño inicial de la evaluación orientada a efectos, contemplaba la participación de un grupo de intervención y un grupo de comparación. Es decir, un grupo de personas que hubieran recibido la intervención, en este caso, del Programa de Atención, y otro grupo con características similares, víctimas de delito, que no hubiese ingresado al Programa. - Sin embargo desde un inicio la contraparte (OAPVD) comunicó que no existía la accesibilidad para disponer de bases de datos de víctimas de delito que no hubieran ingresado al Programa de Atención, ya que éstas listas se encontraban en propiedad del Ministerio Público, siendo datos confidenciales y privados.

				- A partir de esto se determinó trabajar con una muestra para el grupo de comparación de víctimas de delito que habían ingresado al Programa de Atención pero que por alguna razón no continuaron con la intervención, es decir, se definió trabajar con un grupo de comparación no equivalente o contaminado. - El equipo de evaluación, tras los análisis oportunos, determinó que el grupo de comparación podía definirse considerando a las personas usuarias que habiendo ingresado al Programa no habían finalizado su intervención ("falta de interés"). En ese apartado también se explican los criterios de selección de ambas poblaciones (grupo de intervención y grupo de comparación).
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica Los métodos aplicados no cubren aspectos cuanti ni cuali Los métodos aplicados cubren aspectos cuanti o cuali **Los métodos aplicados cubren aspectos cuanti y cuali**	0 1 5 **10**	Se utilizaron tanto métodos de levantamiento de datos cuantitativos (cuestionario) como cualitativos (talleres, entrevistas, observación).
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica Se aplicó sólo un método Se aplicaron dos métodos **Se aplicaron una pluralidad de métodos (más de dos)**	0 1 5 **10**	A partir de la información incluida en el Cuadro nº 7 (p. 62), el alcance de las técnicas utilizadas fue: Técnica / Método | Tipo | Nº: Nº particip. | Perfiles Talleres de trabajo | Presenciales | 8 talleres | Todas las personas funcionarias del Programa de Atención de la OAPVD. Cuestionario | Autoadministrado | 84 | Todas las personas funcionarias de la OAPVD que trabajen actualmente o que hayan trabajado en el Programa de Atención. Entrevistas | Semiestructuradas | 108(grupo interv.) 10(grupo comparac.) | Grupo de intervención: personas usuarias del Programa de Atención. Entrevistas | Semiestructuradas | 26 | Funcionarios/as del Programa de Atención de 3 regiones (Heredia, Central, Limón). Entrevistas | Semiestructuradas | 34 | Stakeholders o actores externos a la OAPVD. Observación | Presenciales | 10 | Oficinas centrales y regionales de la OAPVD. Análisis documental
3.3	**Procedimiento de selección**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección Se explicitan en algunos casos (técnicas) **Se explicitan los procedimientos y criterios de selección**	0 1 5 **10**	Se ha concedido la puntuación máxima, puesto que se constata un esfuerzo importante en tener en cuenta el procedimiento y los criterios de selección en las diversas técnicas, especialmente en algunas como los grupos de intervención y de comparación a los que se realizó entrevistas semiestructuradas (p. 52 Informe final). En otras técnicas también se realizan argumentaciones o motivaciones sobre la idoneidad de los perfiles o de la estrategia de acercamiento a ellos. También es destacable el Cuadro nº 10, que pone de manifiesto los cambios realizados en el diseño evaluativo durante la fase de trabajo de campo, con tres columnas: diseño inicial /limitación / diseño final.
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras?	Ninguna / No aplica **La muestra no es adecuada** La muestra no es adecuada o el error no es aceptable La muestra sí es adecuada y el error aceptable	0 **1** 5 10	El diseño inicial consistía en una muestra estadísticamente representativa (entrevistar a 300 víctimas, 200 de ellas del grupo de intervención, y 100 del grupo de comparación). Sin embargo, fueron numerosas las dificultades para localizar a las víctimas, por tanto el diseño final se tuvo que orientar hacia una muestra no representativa, seleccionada según criterio de experto, con una meta definida por una fecha límite. El cuestionario se repartió a el cual se repartió a todo el personal con experiencia en el Programa de Atención, y se obtuvo una respuesta de 84 cuestionarios auto-administrados realizados. En definitiva, se trata de una muestra que desde el punto de vista de la significatividad estadística no resulta adecuada, si bien teniendo en cuenta las múltiples dificultades logísticas y de despliegue de trabajo de campo, el hecho de conseguir reunir 84 respuestas fue un logro importante.

3.4	Análisis de datos	Calificación (texto)	(numérico)	Observaciones cualitativas
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali Se triangulan datos cuanti-cuali pero no de manera generalizada **Se triangulan datos cuanti-cuali de manera generalizada**	0 1 5 **10**	La triangulación está presente a lo largo del Informe, e impregna todo el proceso de análisis de datos, tanto en resultados como en conclusiones. De hecho, la triangulación aparece mencionada también en el último capítulo sobre lecciones aprendidas y buenas prácticas. En la p.59 del Informe Final se indica lo siguiente: - *"El acercamiento metodológico se basó en un enfoque mixto, es decir, las interrogantes de la evaluación se abordan utilizando tanto técnicas cualitativas como cuantitativas, que permitieron obtener información precisa y analítica sobre los procesos y efectos que ha tenido el Programa de Atención de la OAPVD; llevándose a cabo un proceso de triangulación de la información para el análisis y valoración de las distintas variables evaluadas."* A modo de ejemplo, en una de las conclusiones se señala: - *"De acuerdo con la triangulación de la información recolectada (instrumento a usuarios/as, funcionarios/as y observación), es posible concluir que la infraestructura, en la mayoría de las oficinas visitadas, responde a la prestación de un servicio de calidad. En algunas oficinas, sin embargo, las condiciones no permiten el desarrollo adecuado del trabajo profesional, afectando la calidad de la atención a la persona usuaria". (p. 199).* Y finalmente, en las lecciones y buenas prácticas se pone de manifiesto lo siguiente: - *"Por lo tanto, una buena práctica es la utilización de diferentes técnicas para responder las preguntas de evaluación.* - *Una lección aprendida es que cuando se aplican diferentes técnicas en poblaciones diferentes, se obtiene información valiosa desde lo interno (funcionarios/as en sus diferentes niveles jerárquicos y diversos perfiles profesionales) y externo (personas usuarias y Stakeholders con conocimiento o relación con el programa). De esta manera se incorpora un componente participativo a la evaluación, que sin duda aumentará la calidad".*
3.4.2	¿Se cuantifican las respuestas, siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados….)?	Ninguna / No aplica No se cuantifican las respuestas cuando es posible Se cuantifican las respuestas a veces cuando es posible **Se cuantifican las respuestas de manera generalizada cuando es posible**	0 1 5 **10**	Esta evaluación presta atención a este aspecto, de forma que a la hora de exponer los resultados de la evaluación, y dar respuesta a los diversos interrogantes planteados, generalmente se cuantifican esas respuestas obtenidas. Como se puede ver en los siguientes ejemplos, esto ocurre tanto en los cuestionarios como en técnicas de corte cualitativo, como las entrevistas, si bien hay que señalar que algunas preguntas de esas entrevistas se formulaban en realidad con categorías de respuesta cerradas (ej. grado de acuerdo). Ejemplos: - (Cuestionario) "*Se consultó sobre la percepción del salario recibido según la labor que realizan, un 24,4% califica la remuneración como Muy Buena, un 46,3% como Buena, un 22% como ni mala ni buena y únicamente un 4,9% como mala o muy mala41. En general más del 65% del personal tiene una valoración positiva de la remuneración económica que reciben". (p. 110).* - (Talleres) *"En los talleres de trabajo realizado con las áreas profesionales, se percibió (para todas las disciplinas) una necesidad sentida y expresada por el personal de que su trabajo y esfuerzo sean reconocidos por las jefaturas". (p. 112)* - *(Entrevistas) "Más del 80% de los/as usuarios/as entrevistados/as consideran que el personal fue respetuoso y los trató con dignidad, demostrando interés por el caso y sensibilidad hacia su situación particular". (p. 122).*
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica No se usaron datos del monitoreo **Se usaron datos del monitoreo de manera insuficiente** Se usaron datos del monitoreo de manera suficiente	0 1 **5** 10	*En esta evaluación se realizó un uso profundo y detallado de las bases de datos de personas usuarias, con el objetivo de realizar la muestra para el análisis estadístico. Sin embargo, esta tarea requirió un importante esfuerzo por el equipo evaluador, ya que las bases no eran homogéneas (ej. diferencias en el ordenamiento de los datos, la categorización de las variables, errores detectados con las fechas de ingreso y salida del PA, falta de información, etc.). Toda esta complejidad derivó en que finalmente se tomara la decisión de incluir únicamente el análisis de efectos en las oficinas con mejor calidad de datos.* *Por otro lado, se utilizaron los Informes de rendición de cuentas de OAPVD para ilustrar algunos hallazgos, p.ej. los referidos a recursos financieros.*
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica No se recurrió a fuentes externas **Se recurrió a fuentes externas de manera insuficiente**	0 1 **5**	*Hay dos hechos que influyen en el tipo de datos externos que se incorporan en esta evaluación: se trata de una evaluación que incorpora un notorio componente académico, y el tema tratado es de carácter jurídico.*

		Se recurrió a fuentes externas de manera suficiente	10	*En base a eso, es en el capítulo 2 donde se exponen algunos referentes jurídicos y elementos sobre el contexto institucional (a modo de contextualización, más que formando parte propiamente del análisis o de los resultados de la evaluación).* *Pero en este caso, son más bien fuentes documentales de carácter no estadístico, sino que incluyen contenido del ámbito jurídico-judicial.*
3.5	*Criterios de evaluación*			
3.5.1	*¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa*			
	a) es relevante/pertinente para el país, la región y el grupo meta	**Ninguna / No aplica**	**0**	*La relevancia o pertinencia no es uno de los elementos que guían esta evaluación, ni están incluidos dentro de los interrogantes de evaluación.* *En todo caso, el interrogante que podría estar más próximo a este tipo de cuestiones, aunque de manera indirecta y un tanto tangencial, sería: "¿De qué forma incide la promoción y difusión de la OAPVD en el conocimiento que del Programa y los servicios tienen las personas usuarias y Stakeholders involucrados?".*
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	5	
		Lo expone de forma concluyente y comprensible	10	
	b) logró las metas definidas	**Ninguna / No aplica**	**0**	*En sentido estricto, esta evaluación no se enfoca hacia aspectos de eficacia, sino que se centra en la calidad del servicio, desde una doble perspectiva: evaluación de procesos /formativa, y evaluación de efectos/sumativa.* *En todo caso, los interrogantes de evaluación referidos a efectos específicos, podrían guardar cierta relación con el plano de eficacia o logro de metas definidas, a saber:* *¿De qué forma el Programa de Atención de la OAPVD ha repercutido en el empoderamiento de la víctima durante el proceso penal hasta la sentencia?* *¿De qué manera el Programa de Atención de la OAPVD ha incidido en la re-victimización y la victimización secundaria de las personas atendidas?* *¿De qué manera el programa de atención de la OAPVD ha contribuido a la finalización del proceso penal de las personas atendidas en las oficinas de Heredia, Limón y Central, del 2010 al 2012?*
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	c) es eficiente	Ninguna / No aplica	0	*Uno de los interrogantes de la evaluación, dentro del bloque referido a evaluación de procesos, indica:* *"¿En qué medida los recursos financieros disponibles contribuyen a la prestación de un servicio de calidad del programa de atención?".* *Se realiza un análisis que incluye un balance entre la información financiera disponible en los informes de la OAPVD, y la información cualitativa recopilada mediante entrevistas semi-estructuradas a funcionarios/as de la OAPVD y a Stakeholders. Se incluye, por lo tanto una revisión de:* *Presupuesto neto y deflactado de la OAPVD, y la descripción de su utilización en los últimos años.* *Percepción de la cantidad de recursos financieros por parte del personal y Stakeholders, así como la identificación de necesidades del programa según estos actores.* *Análisis general de la situación de acuerdo a la información recopilada.* *En definitiva, se incluye diversa información que ayuda a entender el papel de los recursos financieros en el conjunto del objeto evaluado, y que guarda relación con aspectos de eficiencia, aunque no se aborde específicamente como tal.*
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	d) causó impactos (outcome e impacto) deseados y no deseados.	Ninguna / No aplica	0	*Los principales efectos deseados que se analizan en esta evaluación son:* *(Adecuada) vivencia del proceso penal de las víctimas* *Empoderamiento de la víctima durante el proceso penal hasta la sentencia.* *Finalización del proceso penal de las personas atendidas en las oficinas analizadas.* *Como principal efecto no deseado, destaca la re-victimización y victimización secundaria de las personas atendidas.* *A la hora de exponer los resultados de la evaluación referidos a estos efectos, se van desgranando los hallazgos obtenidos a través de las diversas técnicas utilizadas. Además, en ocasiones se retoma el marco conceptual de la evaluación, explicado en el anexo 5, para clarificar el significado de cada concepto.* *En definitiva, hay una exposición comprensible y concluyente de los efectos (se usa la terminología de efectos más que la de impactos), tanto deseados como no deseados.*
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	

	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	Esta evaluación no profundiza en la perdubaribilidad de los efectos e impactos, y ello se debe en parte a la propia naturaleza del objeto evaluado (un Programa de Atención a víctimas de delitos), de modo que esa condición de "víctimas "finaliza, al menos en lo formal o jurídico, cuando se da por concluido el proceso penal correspondiente. En definitiva, es un tema sobre el que los interrogantes de la evaluación no entran, y por tanto no se expone información al respecto, lo cual no tiene por qué ser una debilidad, sino que simplemente no procede en este tipo de evaluación.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	Este aspecto no forma parte de los interrogantes de esta evaluación, y por tanto no se abordó en profundidad. Más bien, la información sobre el sistema de monitoreo, centrado sobre todo en las bases de datos disponibles de personas usuarias, se enfocó hacia los aspectos de disponibilidad de información, como fuente de información para la propia evaluación (ej. Para diseñar la muestra de la encuesta).
		No lo expone	**1**	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	Aunque la evaluación no incluye un análisis de impacto propiamente dicho, sino que más bien plantea algunos interrogantes relacionados con el análisis de efectos, puede decirse que los cambios descritos se ponen en relación con el propio objeto evaluado, y de manera suficientemente concluyente. En todo caso, como limitación podría señalarse que la metodología empleada para analizar esos efectos se basa fundamentalmente en elementos perceptivos de la población, a través de la encuesta realizada o de técnicas cualitativas, más que en otro tipo de estrategias de investigación más sofisticadas. El siguiente extracto ejemplifica cómo se atribuye la observación de cambios a la realidad del Programa: - "*El análisis cualitativo muestra que más de la mitad de las personas que se sienten o se han sentido capaces de enfrentar el proceso penal lo atribuyen al apoyo del programa – sin que eso se haya preguntado específicamente. Estas personas resaltan el apoyo emocional recibido del Programa de Atención*". (p. 187)
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	5	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	**10**	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	En el análisis de esta evaluación no se ha dispuesto de TdR, si bien sí se observa que el informe de evaluación responde adecuadamente a los diversos interrogantes de evaluación planteados en el capítulo 3 de diseño evaluativo.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	Aunque el Informe no refleja de manera expresa, diferenciada y claramente reconocible cuáles son los principales puntos fuertes y débiles del objeto evaluado, a lo largo de sus páginas se incluyen tanto aspectos positivos como aspectos negativos del mismo, que requieren mejora. Por tanto, puede decirse que hay una descripción suficientemente equilibrada de ambas cuestiones. Sirva a modo de ejemplo los siguientes aspectos, extraídos del propio Informe: Aspectos positivos: - El Programa ha tenido una incidencia positiva en el empoderamiento de los usuarios/as. - El Programa cuenta con un talento humano profesional, formado y capacitado para cumplir con sus labores. El perfil profesional es acorde con la lógica de intervención planificada. Aspectos negativos: - Las áreas de atención asumen tareas y funciones de otras áreas, lo cual acarrea dificultades. - La población usuaria resaltó que el seguimiento entre cita y cita es algo a mejorar.
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	Una de las debilidades de esta evaluación es que no hay una distinción nítida entre los distintos niveles (hallazgos, conclusiones, recomendaciones), puesto que en ocasiones estos elementos se entremezclan entre sí, creando confusión en el lector, y no permitiendo en consecuencia distinguir de forma fácil y clara entre lo que es descriptivo y lo que es valorativo. 22 Así, al inicio del capítulo 5, titulado "Resultados de la evaluación", se dice que "*dentro de este capítulo se incluyen los hallazgos, conclusiones y recomendaciones según las categorías de análisis incluidas en el proceso de evaluación*". Sin embargo, en el capítulo 6, titulado "Recomendaciones generales", se explica que en él "*se resumen las principales*
		No se distinguen con claridad declaraciones descriptivas y valorativas	**1**	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	10	

				conclusiones y recomendaciones generadas por la evaluación". Al final, las recomendaciones se repiten dos veces, en dos capítulos diferentes. Por tanto, se echa especialmente en falta un capítulo diferenciado sobre conclusiones, que sea sintético en su extensión, y que sirva para tener una visión global y manejable sobre las principales ideas que se derivan de la evaluación realizada. En su lugar, las conclusiones aparecen al final de cada bloque de resultados de evaluación, y resultan demasiado extensas y pormenorizadas, volviendo a recurrirse en ellas a hallazgos concretos.
4.4	¿Se especifica cómo las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica No se especifica **Se especifica en ocasiones** Se especifica en general	0 1 **5** 10	Esta evaluación es bastante extensa, y su capítulo de resultados ocupa más de 100 páginas, lo cual contribuye a que no realice de manera generalizada una mención al grado de fiabilidad asociado a cada hallazgo. No obstante, esta práctica sí se ha identificado en algunas ocasiones, lo que demuestra que esta evaluación presta atención a este aspecto directamente relacionado con la calidad metodológica. Por ejemplo: - "Un obstáculo que se presentó con las personas usuarias del servicio en el año 2010, fue que *un porcentaje significativo no recordaba con detalle el servicio recibido y la vivencia del mismo o tenía muchos problemas para poder diferenciar el servicio del Programa de Atención con la atención recibida de parte del sistema judicial para su caso particular. Además esta población fue la más difícil de localizar (...)Por lo anterior se tomó la decisión de enfocar los esfuerzos restantes para localizar y entrevistar usuarios/as de años posteriores para disminuir dicho sesgo*". (p. 66 y 67) - "*Es importante tomar en cuenta que los resultados de esta entrevista no son representativos de toda la población atendida por el programa y se refieren solamente a la muestra seleccionada*". (171-172) - "*Aunque de estos datos no se puede concluir con seguridad que este empoderamiento es solamente debido al Programa de Atención (o si existen otras variables externas que hayan influido en el empoderamiento), el porcentaje de personas empoderadas o medianamente empoderadas sí indican una influencia positiva del programa*". (p. 172-173). Por otro lado, una de las lecciones aprendidas se refiere a "7.1.7. Evaluabilidad y solicitud de prueba de datos", que entre otros aspectos, aborda la calidad y accesibilidad de los datos.
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica Las conclusiones no están basadas en los resultados Alguna o algunas conclusiones no están basadas en resultados **Todas las conclusiones están basadas en los resultados**	0 1 5 **10**	Los resultados (hallazgos) y las conclusiones aparecen en el mismo capítulo, presentándose las conclusiones después del conjunto de hallazgos expuesto para cada gran bloque de análisis (procesos de atención; talento humano; infraestructura y accesibilidad física; promoción y difusión del programa; recursos financieros; efectos del programa). En definitiva, se observa una adecuada correspondencia entre los resultados obtenidos y las conclusiones planteadas.
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica No se llega ni de forma lógica ni comprensible **Se llega de forma lógica o comprensible pero no las dos** Se llega de forma lógica y comprensible	0 1 **5** 10	El informe no resulta suficientemente claro a la hora de exponer de manera secuencial y diferenciada los hallazgos, las conclusiones y las recomendaciones. De hecho, esos planos se entremezclan en algunas ocasiones, lo cual genera confusión. Y en concreto, las recomendaciones aparecen tanto al final de cada bloque de resultados de la evaluación (una vez expuestos dentro de esos bloques los hallazgos y las conclusiones, aunque tratándose de una extensión tan considerable, no es fácil seguir el hilo conductor que conecta todos esos elementos). Después, vuelve a abordarse las recomendaciones en el capítulo específicamente dedicado a recomendaciones, y que se titula "recomendaciones generales", lo cual conduce a cierta ambigüedad, ya que ahí aparecen también recomendaciones muy concretas. En definitiva, la conexión lógica y comprensible entre conclusiones y recomendaciones resulta insuficiente, y la estructura del Informe no favorece la obtención de esa secuencia por parte del lector.
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica No son relevantes ni específicas **Son relevantes o específicas pero no las dos** Son relevantes y específicas	0 1 **5** 10	El capítulo de recomendaciones tiene una extensión aproximada de 15 páginas, incluyéndose un gran número de recomendaciones, acompañadas de diversas explicaciones o matizaciones. Esto dota al texto de profundidad, pero le resta cierta manejabilidad para su uso en lo operativo. Entre las carencias en el abordaje de las recomendaciones, destacan dos: las recomendaciones aparecen clasificadas por grandes bloques de análisis, pero no priorizadas ni numeradas; no se indica de manera expresa quién debe asumir la responsabilidad de la implementación o el seguimiento de cada recomendación.

				Por otro lado, conviven recomendaciones de diverso alcance y grado de especificidad, por ejemplo: - (Recomendación global): "*Se recomienda analizar que para que la polifuncionalidad sea efectiva, siempre debe de referirse a funciones dentro del área de conocimientos y competencias de las y los profesionales que la asumen. Esto es un punto que se debe de valorar para las tres áreas de atención del Programa*". (p. 196) - (Recomendación muy concreta): "*Un reajuste de los espacios es urgente en aquellas regionales sin privacidad para que el psicólogo dé terapia, y especialmente en aquellas oficinas donde varios profesionales comparten espacio*". (p. 200). En sí mismo, esto no tiene por qué ser una debilidad, pero sí es cierto que resta cierta homogeneidad a este apartado. Por otra parte, quizás hubiera sido deseable una mayor síntesis del contenido de este capítulo, pues es probable que eso potenciara su posterior uso.
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica **El Informe no contiene un resumen** El informe contiene un resumen estructurado El informe contiene un resumen estructurado, bien legible/comprensible	0 **1** 5 10	Una de las principales carencias de esta evaluación es que no dispone de ningún tipo de resumen ejecutivo, que ayude a tener una visión sintética y global sobre los aspectos abordados en la evaluación. Además, esto resultaría especialmente necesario en una evaluación como esta, cuyo Informe Final supera las 200 páginas, y que no ofrece ningún tipo de información sintética, más allá del breve resumen sobre el contenido de cada capítulo que aparece al comienzo, en la Introducción,
5.	**Exhaustividad y adecuación de la documentación**			
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica **Los TdR no son accesibles en anexos** Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción Los TdR son accesibles en la lengua común de trabajo o en una traducción	0 **1** 5 10	Los TdR no se incluyen como material de consulta en los anexos, ni tampoco aparecen en el documento del Informe final.
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica **No se incluyen en anexos** Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	0 **1** 5 10	El documento de anexos reúne aproximadamente 150 páginas, y está formado por 14 apartados o anexos. Sin embargo, en ninguno de ellos se presenta información sobre el cronograma del trabajo de campo, o listados de lugares y organizaciones visitadas o personas entrevistadas. Sí es cierto que en el Informe Final se mencionaba que las actividades del trabajo de campo fueron realizadas entre julio de 2013 y enero del 2014, y también se describen los perfiles entrevistados, o los lugares donde se aplicaron las guías de observación (10 oficinas en diversas regiones del país, llegando a especificarse cuáles). Pero esa información se va incluyendo a lo largo de diversos epígrafes, no siendo una información que aparezca de forma destacada o diferenciada, siendo por tanto difícil localizarlo de forma rápida dentro del Informe o de los Anexos. Por todo ello, se le concede la mínima puntuación.
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica No se incluyen en anexos (guías y fuentes) Se incluyen algunos en anexos (guías o fuentes) **Se incluyen en anexos (guías y fuentes)**	0 1 5 **10**	El Anexo 7, titulado "Instrumentos de recopilación de datos", incluye: - Metodología utilizada en los talleres de trabajo para funcionarios/as del Programa. - Guía de Observación instalaciones OAPVD. - Cuestionario a profesionales. - Cuestionario a técnicos judiciales. - Guiones de entrevista: 1) funcionarios; 2) personas usuarias; 3) Stakeholders externos (alto rango); La Bibliografía aparece, pero no se incluye en los Anexos, sino como último punto del Informe Final.
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos?	Ninguna / No aplica El informe no presenta una estructura adecuada ni elementos que faciliten su lectura **El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo**	0 1 **5**	La claridad y la facilidad de lectura del Informe se ven sacrificada por su extensión y en ocasiones por la densidad o gran cantidad de contenido incluido en cada epígrafe. El Informe incluye algunos de los elementos que tradicionalmente dotan de estructura y ayudan a la lectura, como: índice general, índices específicos (cuadros, gráficos, figuras y flujogramas), resumen del contenido abordado en cada capítulo, realizado en la Introducción.

		ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).		Sin embargo, se echan en falta algunos otros elementos que ayuden a destacar los aspectos más importantes, y que consigan dar una idea resumida de la información obtenida en el proceso de evaluación. En parte, el hecho de tratarse de una evaluación con componente académico ha podido influir en esa prevalencia de la extensión sobre la síntesis. Asimismo, no hay una diferenciación clara y excluyente entre hallazgos, conclusiones y recomendaciones, sino que a veces aparecen ideas repetidas o entremezcladas. En definitiva, esta evaluación obtiene una valoración intermedia en ese ítem, ya que no logra de forma completa aportar una estructura adecuada y una lectura fácil.
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	10	

4.4.9 Evaluación 09: IMAS

0.1 Nombre del proyecto / programa evaluado	**EVALUACIÓN DE EFECTOS GENERADOS POR LOS PROGRAMAS SOCIALES QUE EJECUTA EL IMAS, 2009**
0.2 Institución que implementa (p.ej. Ministerio)	Instituto Mixto de Ayuda Social. IMAS
0.3 Evaluadores	MEPPD Maestría de Evaluación de Programas y Proyectos de Desarrollo. Instituto de Investigaciones en Ciencias Económicas de la Universidad de Costa Rica.
0.4 Monto del proyecto/ programa	¢23.293.163.400 colones (Cuadro 26. Página 70. Apéndice I)
0.5 Monto para la evaluación	¢70.000.000 colones
0.6 Descripción del proyecto/programa evaluado (un párrafo)	
El Instituto Mixto de Ayuda Social (IMAS), es una Institución autónoma con personería jurídica cuya finalidad es resolver el problema de la pobreza extrema en el país, para lo cual planea, dirige, ejecuta y controla planes nacionales destinados a dicho fin. Su objetivo fundamental es el de promover condiciones de vida digna y el desarrollo social de las personas, de las familias y de las comunidades en situación de pobreza o riesgo y vulnerabilidad social, con énfasis en pobreza extrema; proporcionándoles oportunidades, servicios y recursos, a partir del conocimiento de las necesidades reales de la población objetivo, con enfoque de derechos, equidad de género y territorialidad; con la participación activa de diferentes actores sociales y con transparencia, espíritu de servicio y solidaridad. Hay que indicar que esta evaluación de los efectos generados por los programas sociales implementados por el IMAS en el año 2009, se dividió en dos subevaluaciones: ⇨ Evaluación de Efectos del Programa de Asistencia y Promoción Social, año 2009. Con la implementación del Programa de Asistencia y Promoción Social se busca dar apoyo económico para satisfacer necesidades básicas de las familias ubicadas en condiciones de pobreza y vulnerabilidad. El programa está acompañado de acciones de promoción y de orientación individual o familiar que brinden sostenibilidad al impulso inicial, con el objetivo de alcanzar una movilidad social ascendente. ⇨ Evaluación de Efectos del Programa Avancemos, año 2009. El Programa Avancemos se presenta como un programa de transferencia monetaria condicionada, el cual responde a la Rectoría del Sector Social y Lucha contra la Pobreza, con el objetivo de conseguir la Permanencia en el sistema educativo formal de las y los adolescentes y jóvenes entre los 12 y 25 años en condiciones de pobreza extrema, pobreza y vulnerabilidad.	
0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL (máximo media página)	
El trabajo analizado consiste en la evaluación de los resultados generados a partir de la ejecución de los programas sociales, que ejecuta el Instituto Mixto de Ayuda Social IMAS. Atendiendo a la documentación revisada por el equipo que realiza esta meta evaluación no hay indicaciones relativas a la participación de FOCEVAL en el trabajo realizado.	

0.8 Cualificación cuantitativa (con comentarios) según criterios			
78 [Sobre 100]	**10**	***Imparcialidad e independencia:*** [Sobre 10]	La evaluación analizada fue realizada por un equipo técnico externo a la organización que implementa el programa. Asimismo, las fuentes de información seleccionadas garantizaban la inclusión de pluralidad de perspectivas en el proceso evaluativo. Atendiendo a esto, se puede convenir que la evaluación garantizaba su imparcialidad e independencia.
	4,6	***Equidad, participación y transparencia:*** [Sobre 10]	En cuanto a la disponibilidad de documentación hay que poner de manifiesto dos carencias fundamentales: • La documentación final de la evaluación no se encuentra accesible para el público general, al menos vía web. • El equipo que ha realizado la meta evaluación no disponía del APÉNDICE III: LÍNEA BASE ni del APÉNDICE IV: ALGUNOS ASPECTOS METODOLÓGICOS PARA LLEVAR A CABO FUTURAS EVALUACIONES DE EFECTOS. En relación a este aspecto hay que indicar que el Apéndice III era uno de los productos requeridos en los TdR. Respecto a la participación, no se desarrollaron actuaciones tendentes a implicar a los agentes informantes clave de forma previa al lanzamiento de la evaluación.
	40,1	***Credibilidad y exactitud:*** [Sobre 50]	El diseño metodológico de esta evaluación permite atribuirle un importante grado de exactitud y credibilidad. Si bien es cierto que la planificación previa carecía del establecimiento de una cadena de resultados y de hipótesis de impactos, el uso de diferentes técnicas de análisis cuantitativo y cualitativo, la aleatoriedad de la muestra seleccionada y la triangulación de los resultados obtenidos permiten garantizar el desarrollo de un trabajo evaluativo de calidad.
	17,8	***Calidad del informe y utilidad:*** [Sobre 20]	El uso de indicadores en el análisis de la información, la presentación de los resultados finales en base a los elementos que componen los programas, y la profusa redacción de conclusiones y recomendaciones, dotan a la documentación final de unos elevados niveles de calidad y utilidad. Asimismo, hay que indicar que la estructuración de la documentación final de la evaluación en un documento metodológico principal (Informe final) y dos apéndices a modo de "subevaluaciones" de programas (Apéndice I relativo al PAPS y Apéndice II relativo al programa Avancemos) permiten un acceso a la información por parte del lector sencillo, agradable y cobre todo práctico. La falta de un resumen estructurado tanto de la evaluación desarrollada, como de la documentación disponible, es lo que impide alcanzar una nota de 10 en este apartado.

	5,5	***Exhaustividad y adecuación de la documentación:*** [Sobre 10]	Si bien la documentación disponible para el equipo que realiza esta meta evaluación tenía carácter exhaustivo y era adecuada para poder analizarla, se ha revelado como una carencia importante la falta de documentación anexa. En relación a este elemento hay que indicar que no se dispone de información estratégica clave dentro de un proceso de evaluación como puede ser: guías para desarrollar grupos focales, guiones de entrevistas, cronograma de la fase de campo que incluyera listados de lugares visitados y de personas entrevistadas,…

0.9 Resumen de fortalezas de la evaluación.

- ⇨ El equipo evaluador externo ha garantizado la independencia e imparcialidad del trabajo realizado. Al mismo tiempo, se ha incluido una amplia variedad de perspectivas en el desarrollo del análisis cualitativo, incluyendo tanto a población beneficiaria como a gestores y planificadores de los programas evaluados.
- ⇨ El uso de técnicas de análisis cuantitativo que contemplan la selección aleatoria de participantes supone contar con la posibilidad de poder inferir los resultados obtenidos con la muestra elegida a la población total beneficiaria. En relación a este aspecto hay que indicar que la muestra poblacional seleccionada destaca por su amplitud y nivel de confianza (más de 2.000 familias beneficiarias analizadas).
- ⇨ El equipo evaluador ha desarrollado un trabajo de triangulación más que aceptable, obteniendo información de carácter estratégico tanto del ámbito cualitativo como cuantitativo. Al mismo tiempo, la triangulación ha permitido cotejar información obtenida a partir de distintas fuentes, pero relativa a un mismo elemento evaluado.
- ⇨ El posicionamiento con el que se ha realizado el análisis de los resultados obtenidos ha estado focalizado a la consecución de una valoración de los efectos generados por los programas sociales de IMAS evaluados (PAPS y Programa Avancemos) en la población beneficiaria. Al mismo tiempo, la planificación del proceso evaluativo y su diseño metodológico, han posibilitado la obtención de resultados vinculados a la mejora en la gestión de ambos programas, a su eficacia y a su eficiencia.
- ⇨ Tanto la decisión de estructurar la documentación final en varios elementos (Informe final, Apéndice I y Apéndice II) como su redacción, dotan a la documentación disponible de esta evaluación de un elevado nivel de calidad y utilidad. Asimismo, este trabajo de evaluación ha respondido adecuadamente a los requerimientos incluidos en los TdR.

0.10 Resumen de debilidades de la evaluación.

- ⇨ Atendiendo a la situación socio económico de la población beneficiaria de los programas sociales evaluados, se hacía necesario establecer un proceso previo de sensibilización/presentación que facilitara la participación de la muestra seleccionada.
- ⇨ Durante el desarrollo de esta evaluación se produjo una situación conflictiva entre el diseño teórico inicial de la evaluación y la realidad encontrada en el momento de ejecución de la misma. En este sentido hay que poner de manifiesto que debido a elementos ajenos al equipo evaluador (imposibilidad para localizar a población seleccionada en la muestra, falta de participación de los beneficiarios…) hubo que acometer revisiones metodológicas en el proceso evaluativo tales como cambios en la muestra, manteniendo su carácter aleatorio pero ampliando la selección inicial.
- ⇨ Se registraron incidencias en las consultas realizadas a la base de datos SIPO. En la mayor parte de los casos se trataba de diferencias existentes entre la información digitalizada y la física a nivel de número de expediente.
- ⇨ El hecho de efectuar un proceso de selección aleatoria de participantes, disponiendo de una muestra poblacional bastante amplia, y aplicando unos criterios técnicos que minimizan el error permitido y elevan el nivel de confianza de los análisis cuantitativos, permite desarrollar un análisis exhaustivo que posibilita la inferencia de resultados muestrales a la población general. Sin embargo, aun siendo esto así, la no existencia de un grupo de control impide en modo alguno poder inferir los efectos detectados por la evaluación a la exclusiva aplicación de los programas sociales del IMAS analizados.
- ⇨ Existe una carencia importante en relación a la documentación disponible en tanto que no se incluye el Apéndice III relativo al establecimiento de la línea de base para el año 2009. La importancia de este documento era bastante elevada ya que era uno de los productos requeridos en los TdR.

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1		Ninguna / No aplica	0	

	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Evaluación interna (100% trabaja en instituciones evaluadas)	1	Se trata de una evaluación externa en tanto que el equipo evaluador pertenece a la Maestría de Evaluación de Programas y Proyectos de Desarrollo de la Universidad de Costa Rica, entidad independiente del Instituto Mixto de Ayuda Social IMAS. Sin embargo, aunque se trata de una evaluación externa, hay que indicar que, según información contenida en los TdR, el cliente (IMAS) aportó sus instalaciones al tiempo que suministró cuanta información (expedientes de beneficiarios) solicitó el equipo evaluador.
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	5	
		Evaluación externa (100% externo a instituciones evaluadas)	**10**	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y análisis de datos y presentación de resultados	Ninguna / No aplica	0	Las fuentes de información elegidas por el equipo evaluador se definieron desde una doble perspectiva: • Familias beneficiarias de programas sociales del IMAS (beneficiarios directos). • Funcionarios y funcionarias, gerentes regionales y responsables de unidades de gestión de IMAS (planificadores y gestores del programa).
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	5	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos).	**10**	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos? **Fases**: planificación e implementación.	Ninguna / No aplica / No hay información en los documentos consultados	0	En base a la información disponible en los documentos finales de evaluación revisados hay que indicar que no se involucró a grupos interesados en la planificación e implementación del proceso de evaluación. La planificación e implementación fue desarrollada por el equipo evaluador en base a los requerimientos establecidos por el cliente (IMAS)
		No hay participación	**1**	
		Participan escasa o desigualmente en planificación o implementación	5	
		Participan activamente en la planificación y en la implementación	10	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	El documento que contiene el Informe final y documento relativo al Apéndice I "Evaluación de Efectos del Programa de Asistencia y Promoción Social" no incluyen preguntas de evaluación. El Apéndice II "Evaluación de Efectos del Programa Avancemos" incluye un amplio listado de preguntas de evaluación en su epígrafe 3.3 Interrogantes de evaluación, redactadas de forma clara, y adaptadas al contexto del objeto de evaluación.
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	5	
		Las preguntas son claras y adaptadas	**10**	
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	Ninguna / No aplica	0	La redacción de los TdR por parte del organismo contratante es clara y comprensible. Se incluyen de forma detallada, al tiempo que estructurada, todos los elementos necesarios para presentar una oferta completa que atienda los requerimientos necesarios del trabajo evaluativo solicitado.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	**10**	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	Ni en el diseño de la evaluación, ni en su implementación, se contempló la realización de un taller inicial con stakeholders que facilitaran su participación. En relación a este aspecto, hay que indicar que la coordinación del trabajo de campo estuvo a cargo del "Proyecto Aplicación, Digitación y Revisión de Fichas de Información Social del IMAS", instalado en la Universidad de Costa Rica. Esta entidad universitaria cuenta con una dilatada experiencia en la visitación domiciliaria de las familias beneficiarias del IMAS.
		No hubo un taller de inicio con stakeholders para facilitar su participación	**1**	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	
		Hubo un taller de inicio con stakeholders	10	
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica	0	No existe documentación accesible para el público relativa a este trabajo de evaluación.
		No son accesibles para el público (ni Informe ni resumen)	**1**	

		Son accesibles para el público (Resumen o extracto)	5	
		Son accesibles para el público (Informe y resumen)	10	
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	El objetivo de la evaluación se describe en el epígrafe 1.2 del documento relativo al Informe final, y obedece en todo momento a los requerimientos incluidos en los TdR. En relación a este aspecto hay que indicar que, si bien los objetivos se establecen en el documento del Informe final, cada uno de los Apéndices (I-II) incluye una descripción concreta del objeto de evaluación.
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	
		Se describe el objetivo de la evaluación de manera imprecisa	5	
		Se describe de manera precisa el objetivo	**10**	
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica	0	El objeto de la evaluación es descrito de forma concreta en los dos documentos Apéndices de la evaluación, estando cada uno vinculado al programa evaluado (Apéndice I Programa PAPS, Apéndice II Programa Avancemos). Si bien en el epígrafe relativo al objeto de la evaluación (3.1 en ambos documentos) no se detalla suficiente información sobre el mismo, en el capítulo previo (2 en ambos documentos) se definen los programas evaluados indicando información relativa al marco legal, población beneficiaria, información financiera, planificación temporal y regional,…
		No se describe el objeto de evaluación	1	
		Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes	5	
		Se describe con precisión (al menos los ítems contemplados)	**10**	
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica	0	El proceso de selección de la muestra de participantes en la evaluación se realizó de forma aleatoria. Sin embargo, hay que indicar varios aspectos que pueden suponer un análisis previo de los stakeholders: • Se incluyó de forma específica en las entrevistas grupales a población afectada por el terremoto de Cinchona, por haberse producido la emergencia el año de estudio. • En base a su conocimiento experto, las Direcciones Regionales seleccionaron, a partir de ciertos lineamientos, a los participantes en las entrevistas grupales.
		No se realizó un análisis de stakeholders	**1**	
		Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.)	5	
		Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	10	
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica	0	Ninguno de los documentos finales analizados (Informe final, Apéndice I y Apéndice II) incluye información relativa a una teoría de la intervención, cadena de resultados o lógica de impacto, asociada a los programas evaluados.
		No se elaboró	**1**	
		Se elaboró pero no se usa en la evaluación o no es concluyente	5	
		Se elaboró, se usa en la evaluación y es concluyente	10	
	¿Se formularon hipótesis de impacto concluyentes?	Ninguna / No aplica	0	Ninguno de los documentos finales analizados (Informe final, Apéndice I y Apéndice II) establece hipótesis de impacto concluyentes.
		No se formularon	**1**	
		Se formularon pero no son concluyentes	5	
		Se formularon y son concluyentes	10	
	¿Se diferencia entre los diferentes niveles de impacto? Niveles de impacto:	Ninguna / No aplica	0	En el caso de la evaluación de Efectos del PAPS (Apéndice I) se analizó la información cuantitativa diferenciando entre distintos tipo de beneficios del programa (Cuadro 2, página 26), mientras que la cualitativa se analizó en base a diferentes dimensiones y categorías (Recuadro 9, página 32). En relación a la evaluación de Efectos del programa Avancemos (Apéndice II) la información cuantitativa se estudió diferenciando entre objetivos del Programa, población objetivo y criterios de selección, y condicionalidad que deben de cumplir los beneficiarios (Página 20). La información cualitativa se evaluó en base a diferentes temas y categorías (Recuadro2, página 27).
		No se diferencia entre niveles de impacto	1	
		Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)	**5**	
		Se diferencia con claridad los diferentes niveles de impacto	10	
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica	0	Atendiendo a la información contenida en los documentos finales de evaluación, se puede afirmar que se ha descrito de forma exhaustiva el contexto de la evaluación, relacionándolo con la teoría de intervención del proyecto. En relación a este aspecto hay que indicar que, mientras el Informe final contiene una información más global del contexto y de la teoría de la intervención, en los apéndices de los proyectos evaluados (Apéndice I PAPS y Apéndice II Avancemos) se detalla la información concreta relativa a cada proyecto a nivel de contexto.
		No se describe el contexto de la evaluación	1	
		Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención	5	
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	**10**	
3.2	**Diseño y métodos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**

<table>
<tr><td>3.2.1</td><td>¿Se corresponde el diseño de la evaluación con los intereses de la misma?</td><td>Ninguna / No aplica (Los TdR no explicitan los intereses de la evaluación)
El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)
El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)
Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)</td><td>0
1
5
10</td><td>El diseño metodológico de esta evaluación se corresponde de una forma más que adecuada con los requisitos establecidos para la misma que se detallan en el documento de los TdR.</td></tr>
<tr><td>3.2.2</td><td>¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?</td><td>Ninguna / No aplica
No se explicitan ni justifican el diseño y los métodos aplicados
Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados
Se explicitan y justifican suficientemente el diseño y los métodos aplicados</td><td>0
1
5
10</td><td>El diseño metodológico de la evaluación y los métodos aplicados en la misma se detallan y justifican de forma concreta en la documentación final entregada.
En relación a este aspecto hay que indicar que la explicación y justificación metodológica se hace de forma separada. Así, mientras en el documento del Informe final se incluye el diseño metodológico utilizado (Capítulo II. Aspectos metodológicos) en cada uno de los apéndices se incluye el modelo de evaluación y procedimientos empleados para cada programa (Capítulo IV de ambos apéndices).</td></tr>
<tr><td>3.2.3</td><td>¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)</td><td>Ninguna / No aplica
No se consideró lo contrafactual
Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró
Se consideró lo contrafactual y se especificó cómo</td><td>0
1
5
10</td><td>La evaluación analizada no consideró el uso de lo contrafactual.
Tal y como se indica en el Informe final (Epígrafe 2.1.1 Sobre el Diseño Muestral del Programa I y el Programa II) “por consideraciones propias de los Programas, se contempló llevar a cabo la evaluación de efectos del año 2009, sin grupo de control.”</td></tr>
<tr><td>3.2.4</td><td>¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?</td><td>Ninguna / No aplica
Los métodos aplicados no cubren aspectos cuanti ni cuali
Los métodos aplicados cubren aspectos cuanti o cuali
Los métodos aplicados cubren aspectos cuanti y cuali</td><td>0
1
5
10</td><td>El trabajo de evaluación realizado contempló la utilización de métodos de levantamiento de datos que cubrieron aspectos cuantitativos (encuestas) y cualitativos (entrevistas personales).</td></tr>
<tr><td>3.2.5</td><td>¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)?
a) Entrevistas guiadas (número?) ; b) grupos focales (número?)
c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...</td><td>Ninguna / No aplica
Se aplicó sólo un método
Se aplicaron dos métodos
Se aplicaron una pluralidad de métodos (más de dos)</td><td>0
1
5
10</td><td>Apéndice I Evaluación de Efectos del Programa de Asistencia y Promoción Social PAPS
<table>
<tr><th rowspan="2">Técnica / Método</th><th rowspan="2">Tipo</th><th colspan="2">Nº</th></tr>
<tr><th>Ediciones</th><th>Participantes</th></tr>
<tr><td>Entrevistas</td><td>guiadas</td><td>4</td><td>23</td></tr>
<tr><td>Entrevistas</td><td>grupales</td><td>10</td><td>88</td></tr>
<tr><td>Encuestas</td><td>Face to face, telefónicas, online</td><td>1</td><td>1.397</td></tr>
</table>
Apéndice II Evaluación de Efectos del Programa Avancemos
<table>
<tr><th rowspan="2">Técnica / Método</th><th rowspan="2">Tipo</th><th colspan="2">Nº</th></tr>
<tr><th>Ediciones</th><th>Participantes</th></tr>
<tr><td>Entrevistas</td><td>guiadas</td><td>4</td><td>23</td></tr>
<tr><td>Entrevistas</td><td>grupales</td><td>10</td><td>78</td></tr>
<tr><td>Encuestas</td><td>Face to face, telefónicas, online</td><td>1</td><td>1.269</td></tr>
</table></td></tr>
</table>

				Las entrevistas guiadas fueron comunes a la evaluación de ambos programas (PAPS y Avancemos). Su distribución fue la siguiente: • Entrevistas con funcionarias y funcionarios (7) • Entrevistas no estructuradas autoadministradas remitidas a las/los Gerentes Regionales (9) • Entrevista no estructurada autoadministrada remitida a la Unidad Coordinadora del Programa Avancemos (1) • Entrevistas con actores relevantes (6)
3.3	**Procedimiento de selección**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica No se explicitan los procedimientos ni criterios de selección Se explicitan en algunos casos (técnicas) **Se explicitan los procedimientos y criterios de selección**	0 1 5 **10**	El proceso de selección y los criterios aplicados para el mismo se detallan en el documento que contiene el Informe final. En relación a esto hay que indicar que en el caso de la información cuantitativa, el proceso de selección de la muestra se detalla de forma explícita en el epígrafe 2.1.1 Sobre el Diseño Muestral del Programa I y el Programa II. En el caso de la selección de los agentes informantes clave de aspectos cualitativos, el procedimiento se describe en el epígrafe 2.1.6 Metodología sobre otras técnicas e instrumentos empleados.
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? **Se entiende que aplica esta pregunta para las Encuestas**	Ninguna / No aplica La muestra no es adecuada La muestra no es adecuada o el error no es aceptable **La muestra sí es adecuada y el error aceptable**	0 1 5 **10**	Se definió una muestra inicial de 2.400 familias, seleccionadas de forma aleatoria y estratificada para 1.200 familias del Programa Avancemos y 1.200 familias del PAPS. En referencia a la calidad de la muestra extraída, los márgenes de error máximo permitidos están por debajo del 5%, presentando un nivel de confianza del 95%, según el universo muestral que fue de 152.966 personas beneficiadas.
3.4	**Análisis de datos**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica No se triangulan datos cuanti-cuali Se triangulan datos cuanti-cuali pero no de manera generalizada **Se triangulan datos cuanti-cuali de manera generalizada**	0 1 5 **10**	La triangulación de datos cuantitativos y cualitativos es utilizada en la evaluación de ambos programas (PAPS y Avancemos). En ambos Apéndices se presentan los criterios concluyentes de la evaluación sostenidos en parte por el cotejo de la información cuantitativa con la cualitativa Asimismo, en el análisis cualitativo se efectuó una triangulación de la información por técnica utilizada y por actores participantes.
3.4.2	¿Se cuantifican la respuesta siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados…?	Ninguna / No aplica No se cuantifican las respuestas cuando es posible Se cuantifican las respuestas a veces cuando es posible **Se cuantifican las respuestas de manera generalizada cuando es posible**	0 1 5 **10**	El análisis de datos efectuado en ambos Apéndices presenta la información de manera cuantificada. Esto es así tanto para el análisis de información (cuantitativa y cualitativa) como para la presentación de resultados y criterios concluyentes.
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica No se usaron datos del monitoreo Se usaron datos del monitoreo de manera insuficiente **Se usaron datos del monitoreo de manera suficiente**	0 1 5 **10**	En ambos casos (Evaluación del PAPS y Evaluación del programa Avancemos) se usaron datos extraídos de los sistemas de monitoreo del IMAS. En relación a esta cuestión hay que indicar que uno de los encargos de este trabajo de evaluación fue el de realizar la revisión de los expedientes de las familias beneficiarias seleccionadas en la muestra.
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica **No se recurrió a fuentes externas** Se recurrió a fuentes externas de manera insuficiente Se recurrió a fuentes externas de manera suficiente	0 **1** 5 10	En base a la información presentada en la documentación final de este trabajo de evaluación, se puede afirmar que los datos utilizados para el análisis se extrajeron de fuentes internas. Las fuentes de datos utilizadas estaban contenidas en dos registros: • SABEN Sistema de registro de beneficiarios del IMAS. • SIPO Sistema de Información de Población Objetivo del IMAS.
3.5	**Criterios de evaluación**	**Calificación (texto)**	*(numérico)*	**Observaciones cualitativas**
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			

	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	En el documento que incluye el Informe final se expone de manera somera la relevancia y pertinencia del IMAS y de los proyectos sociales que implementa. En el caso de los Apéndices que presentan las Evaluaciones específicas de cada programa (PAPS y Avancemos) se detalla de forma profusa, clara y detallada la importancia de cada programa para la población. El detalle de cada proyecto se presenta en el Capítulo II de cada Apéndice.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	5	
		Lo expone de forma concluyente y comprensible	**10**	
	b) logró las metas definidas	Ninguna / No aplica	0	Tanto desde un punto de vista cualitativo como cuantitativo, ambos Apéndices presentan los resultados y efectos que, desde un punto de vista cuantitativo y cualitativo, han tenido los programas sociales evaluados (PAPS Y Avancemos) sobre la población beneficiaria. Asimismo, hay que indicar que en ambos documentos se exponen los efectos derivados de la aplicación de los programas implementados por el IMAS de forma concluyente y comprensible.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	c) es eficiente	Ninguna / No aplica	0	El trabajo de evaluación realizado expone de forma concluyente y compresible el análisis de la eficiencia de los programas evaluados. En relación a este aspecto hay que indicar que en ambos Apéndices se encuentran apartados específicos relativos a la eficiencia del programa evaluado en cada caso. En del Apéndice I (PAPS) dentro del apartado relativo a la presentación de Criterios concluyentes, se incluye el epígrafe 6.1.4 Valoración de la percepción de la eficiencia situación familiar. En el caso del Apéndice II (Programa Avancemos) se incluye el estudio de la eficiencia dentro del Análisis de la información, epígrafe 4.4.3.2.6 Valoración de la Eficiencia, y también dentro de la Presentación de los resultados, epígrafe 5.3.2.6 Valoración de la eficiencia.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	d) causó impactos (outcome e impacto) deseados y no deseados.	Ninguna / No aplica	0	Analizando este trabajo de evaluación hay que poner de manifiesto que, atendiendo a toda la documentación disponible, en ningún momento se contempla la estimación de resultados e impactos de los programas evaluados. Sin embargo, hay que indicar que el elemento central de la evaluación era el de estimar efectos derivados de la implementación de programas sociales por parte del IMAS. En ese sentido, hay que poner de manifiesto que la exposición de los efectos derivados de los programas evaluados se realizó de forma concluyente y comprensible, incluyéndose esta información en los Apéndices I y II, vinculada en cada caso al programa evaluado (PAPS y Programa Avancemos). Se ha optado por interpretar el estudio de los efectos como una adecuada a esta pregunta
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	Ídem del anterior. Valga la observación cualitativa anterior para este apartado en tanto que el objeto del mismo sigue siendo la estimación de impactos, aunque en este caso desde una perspectiva temporal.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	El Instituto Mixto de Ayuda Social cuenta con dos registros fundamentales para la gestión de programas sociales. Por una lado se encuentra Sistema de registro de beneficiarios del IMAS (SABEN) y por otro el Sistema de Información de Población Objetivo del IMAS (SIPO). Ambos registros permiten efectuar tanto el monitoreo de la gestión de programas, como proporcionar información fundamental para implementar procesos evaluativos.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	Atendiendo a la información contenida en la documentación final de la evaluación se puede convenir que los cambios identificados en la población beneficiaria se vinculan con la participación en el programa de manera concluyente. Esta afirmación se basa en tres aspectos clave: • La selección de la muestra objeto de estudio se hizo de forma aleatoria, metodología que permite realizar inferencias válidas a la población objeto de estudio. • Los participantes en el trabajo de campo relativo al análisis cualitativo fueron seleccionados a partir de la muestra aleatoria anterior, o incluidos por su posición estratégica dentro de la gestión del programa. • Se efectuó una triangulación que sirvió de cotejo entre los resultados cuantitativos y los cualitativos.
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	5	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	**10**	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas

4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	La documentación final de la evaluación (Informe final, Apéndice I y Apéndice II) responde perfectamente al listado de preguntas establecidas en los TdR, y obedece en todo momento a los requerimientos efectuados en los mismos.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	Las evaluaciones presentadas en los apéndices (Apéndice I Evaluación de Efectos del PAPS y Apéndice II Evaluación de Efectos del programa Avancemos) describen a través de criterios concluyentes los aspectos fuertes y débiles de los programas evaluados.
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	Tanto en el caso del Apéndice I como en el del Apéndice II, la información se presenta de forma que resulta sencillo distinguir entre reflexiones descriptivas y valorativas. En este sentido se pone de manifiesto claramente el trabajo de triangulación realizado por el equipo evaluador que suponía, en parte, cotejar la información cualitativa obtenida con los resultados del análisis cuantitativo.
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	**10**	
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	Las fuentes de información cualitativa son de carácter primario (familias beneficiarias, gestores y planificadores de los programas). Asimismo, atendiendo a la caracterización de los participantes seleccionados es confiable. Las fuentes de información cuantitativa son registros del IMAS de carácter primario y confiable: • SABEN Sistema de registro de beneficiarios del IMAS. • SIPO Sistema de Información de Población Objetivo del IMAS.
		No se especifica	1	
		Se especifica en ocasiones	5	
		Se especifica en general	**10**	
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica	0	Las conclusiones de la evaluación realizada se presentan de forma separada en cada uno de los Apéndices (I y II), y se vinculan al programa evaluado (PAPS y Avancemos) que corresponda según el documento. En este sentido, las conclusiones de la evaluación de los efectos de cada uno de los programas estudiados se basan en los resultados que se obtuvieron a partir del análisis de la información recabada. Se trata de un listado de conclusiones fundamentadas en el análisis cuantitativo, cualitativo y en la triangulación de diferentes fuentes de información.
		Las conclusiones no están basadas en los resultados	1	
		Alguna o algunas conclusiones no están basadas en resultados	5	
		Todas las conclusiones están basadas en los resultados	**10**	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	Los apéndices de las evaluaciones de ambos programas (PAPS y Programa Avancemos) presentan los resultados, conclusiones y recomendaciones de forma estructurada, lógica y comprensible. En ambos documentos finales de evaluación la estructura sigue la siguiente secuencia: Resultados > Criterios concluyentes > Conclusiones > Recomendaciones
		No se llega ni de forma lógica ni comprensible	1	
		Se llega de forma lógica o comprensible pero no las dos	5	
		Se llega de forma lógica y comprensible	**10**	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	El listado de recomendaciones propuesto es relevante y específico para la mejora del programa, teniendo en cuenta siempre a la población beneficiaria. Se trata en ambos casos de un amplio listado de propuestas (51 recomendaciones para la mejora del PAPS y 25 para Avancemos). En relación a este ítem hay que indicar que en el Apéndice I (Evaluación de efectos del PAPS) las recomendaciones se estructuran en base los distintos elementos del programa: Atención a necesidades básicas, Emergencias, Capacitación, Veda, Ideas productivas, Fideicomiso y Mejoramiento de la vivienda.
		No son relevantes ni específicas	1	
		Son relevantes o específicas pero no las dos	5	
		Son relevantes y específicas	**8-9-10**	
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	La documentación final presentada de la que dispone el equipo que realiza esta meta evaluación (Informe final, Apéndices I y II) no incluye un resumen estructurado de la evaluación realizada. Si bien es cierto que la articulación de un trabajo de tan amplia magnitud en varios documentos es adecuada y dota al trabajo de un mayor nivel de utilidad, se echa de menos un documento que a modo de resumen presente las claves esenciales del trabajo realizado de forma conjunta.
		El Informe no contiene un resumen	**1**	
		El informe contiene un resumen estructurado	5	
		El informe contiene un resumen estructurado, bien legible/comprensible	8-9-10	
5.	**Exhaustividad y adecuación de la documentación**			

5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica	0	Si bien los TdR no se encuentran accesibles en forma de Anexo dentro de los tres documentos finales de la evaluación, sí se incluían como documento independiente (formato Word) dentro del conjunto de archivos recibidos por el equipo que realiza esta meta evaluación.
		Los TdR no son accesibles en anexos	1	
		Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	
		Los TdR son accesibles en la lengua común de trabajo o en una traducción	**10**	
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	En el caso de los documentos relativos a los Apéndices 1 y 2, no se incluyen Anexos. En relación al Informe final, si bien se indica en su índice la existencia de tres anexos finales, incluyendo uno con las "Guías de las entrevistas grupales no estructuradas" (Anexo 3), el documento revisado tan sólo incluye el Anexo 1 relativo a "Análisis de sensibilidad: Muestra versus aplicación ficha de información social (FISEVA), programa Avancemos". La información existente relativa a las entrevistas se encuentra en los epígrafes 2.1.4 Sobre las Entrevistas Grupales con Beneficiarios y 2.1.6 Metodología sobre otras técnicas e instrumentos empleados, sin embargo no incluyen un listado de lugares y organizaciones visitadas y de personas entrevistadas.
		No se incluyen en anexos	**1**	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	5	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	En el caso de los documentos relativos a los Apéndices 1 y 2, no se incluyen Anexos. En relación al Informe final, si bien se indica en su índice la existencia de tres anexos finales, incluyendo uno con las "Guías de las entrevistas grupales no estructuradas" (Anexo 3), el documento revisado tan sólo incluye el Anexo 1 relativo a "Análisis de sensibilidad: Muestra versus aplicación ficha de información social (FISEVA), programa Avancemos". La información existente relativa a las entrevistas se encuentra en los epígrafes 2.1.4 Sobre las Entrevistas Grupales con Beneficiarios y 2.1.6 Metodología sobre otras técnicas e instrumentos empleados, sin embargo no incluyen guías. En referencia a la literatura y documentación utilizada, en el capítulo 5 del Informe final se incluye un listado de la Bibliografía utilizada.
		No se incluyen en anexos (guías y fuentes)	**1**	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	10	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	Ninguna / No aplica	0	A excepción de los resúmenes de los capítulos, los tres documentos finales de evaluación presentan una estructura más que adecuada para facilitar su lectura. Tanto el Informe final como los Apéndices 1 y 2, presentan una jerarquización a nivel de capítulos, y una distribución de epígrafes dentro de los mismos, que facilitan su lectura y entendimiento. Asimismo, los tres documentos contienen gráficos y tablas en un número más que suficiente, presentando siempre la información de forma clara, concisa y detallada.
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).	5	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	**10**	

4.4.10 Evaluación 010: PRENATAL

0.1 Nombre del proyecto / programa evaluado	**ATENCIÓN PRENATAL EN EL PRIMER NIVEL DE ATENCIÓN EN LA REGIÓN BRUNCA Y HUETAR CARIBE**
0.2 Institución que implementa (p.ej. Ministerio)	Ministerio de Salud
0.3 Evaluadores	Ministerio de Salud Caja Costarricense de Seguro Social (CCSS) Asesoría técnica: Ministerio de Planificación Nacional y Política Económica (MIDEPLAN); GIZ-FOCEVAL (asesor etapa de programación)
0.4 Monto del proyecto/ programa	No especificado en el Informe final de evaluación.
0.5 Monto para la evaluación	No especificado en el Informe final de evaluación.

0.6 Descripción del proyecto/programa evaluado (un párrafo)
El programa evaluado es el de Atención Prenatal en el Primer Nivel de Atención, que se realizó en Costa Rica en la Región Brunca y la Región Huetar Caribe en el año 2012. Se trata de un programa de atención prenatal que se brinda en los Equipos Básicos de Atención Integral en Salud (EBAIS) vinculado al cumplimiento de la Norma Oficial Atención Integral a la Mujer Durante el Embarazo de Bajo Riesgo Obstétrico. El programa se inserta dentro del Plan Estratégico Nacional para una Maternidad e Infancia Saludable y Segura, 2006-2015 cuyo propósito es "garantizar la salud y el bienestar de las mujeres, las madres, las niñas y los niños, para que tengan oportunidad de una vida saludable y segura, mediante el acceso oportuno y de calidad a salud, educación, trabajo, transporte, en un entorno familiar y social solidario, amigable y seguro". Este Plan contempla los siguientes componentes: 1) atención integral en salud materna, perinatal e infantil, 2) aseguramiento de la calidad, 3) participación y comunicación social, y 4) marco de política pública legal e institucional.
0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL (máximo media página)
Para el desarrollo de esta evaluación se contó con el Modelo Conceptual y Estratégico de Evaluación de Impacto de Acciones en Salud diseñado por la Unidad de Evaluación de Impacto de Acciones en Salud (UEIAS) del Ministerio de Salud, el cual se apunta que está alineado con el Manual Gerencial para el Diseño y Ejecución de Evaluaciones Estratégicas de Gobierno (MIDEPLAN). Partiendo de que al Ministerio de Salud como ente rector de la Producción Social de la Salud, le compete desarrollar la función estratégica de Evaluación del Impacto de las Acciones en Salud las autoridades plantearon la necesidad de desarrollar esta evaluación a través de un proceso de coordinación y articulación de la Unidad de Evaluación de Impacto de las Acciones en Salud de la Dirección de Planificación Estratégica de la Salud (Ministerio de Salud), con la Agencia de Cooperación Alemana (GIZ), el Programa de Fomento de las Capacidades en Evaluación (FOCEVAL) y el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN). Según se apunta en el informe "el representante de FOCEVAL participó hasta la etapa del inicio de la elaboración del diseño de investigación" (Pág.18). "La formulación del diseño de evaluación se dio en el marco de un acuerdo de colaboración entre Ministerio de Salud, MIDEPLAN y FOCEVAL, en una primera etapa. Posteriormente, se integró a la CCSS, como prestador de servicios de salud, conformándose un equipo de trabajo altamente participativo e intersectorial, que continuó en el desarrollando del proceso. Uno de los aspectos prioritarios para la implementación de este estudio fue el diseño de la Cadena de Resultados". (Pág.22)

0.8 Cualificación cuantitativa (con comentarios) según criterios

58,1 [Sobre 100]	**5**	***Imparcialidad e independencia:*** [Sobre 10]	La selección del equipo evaluador y de los stackeholders participantes, no garantiza en este trabajo unos niveles elevados de imparcialidad e independencia. En relación a esto hay que indicar que, si bien inicialmente se contó con la participación de una importante diversidad de actores, tanto durante la realización del trabajo de campo, como en el análisis de información y resultados, el trabajo sólo incluyó a agentes pertenecientes al programa evaluado (técnicos del ministerio de salud, gestores del programa y usuarias del servicio).
	7,8	***Equidad, participación y transparencia:*** [Sobre 50]	Aunque en la planificación inicial se contempló la participación de grupos de diferentes entidades que aportaron diversidad de perspectivas, esta situación no tuvo continuidad durante los trabajos de evaluación. No existe disponibilidad de toda la documentación. En relación a este aspecto hay que indicar que el informe final no se encuentra accesible para el público general, al menos vía web.
	32,6	***Credibilidad y exactitud:*** [Sobre 20]	A partir del análisis de la información presentada en el Informe final se puede afirmar que el documento presenta unos niveles de credibilidad y exactitud mínimos. En relación a esto, se detallan a continuación varios aspectos que justifican la afirmación realizada: • Falta información relativa a aspectos económicos del programa y de la propia evaluación. • No se desarrolla un análisis previo de los agentes informantes clave. • Aunque en el documento se indica que se va a usar un "Modelo Conceptual y Estratégico de Evaluación de Impacto de Acciones en Salud", el diseño metodológico no incluye elementos esenciales en una evaluación de impacto como puede ser la definición de un contra-factual, el establecimiento de hipótesis previas para dar respuesta a las preguntas de evaluación, o la determinación de diferentes niveles de impacto. • No se efectúa un proceso de triangulación de la información que permita validar los hallazgos registrados en el trabajo de campo. • El informe final no expone de manera concluyente y compresible los logros alcanzados por el programa.

	7,5	***Calidad del informe y utilidad:*** [Sobre 10]	En base a la información presentada en el Documento final, los niveles de calidad del informe son bajos. Se trata de una evaluación de la calidad de un servicio público que se ha condicionado previamente con la aplicación metodológica de un modelo de evaluación de impacto. Sin embargo, tanto la definición del diseño metodológico como el desarrollo del trabajo de campo no han cumplido con los requisitos mínimos exigibles para una evaluación de impacto. Asimismo, aunque en el Informe final se parte de una premisa que vincula directamente un nivel de cuidados prenatales aceptables con la aparición de al menos cuatro resultados positivos, no se identifican ni resultados ni impactos del programa. Se trata por tanto de un documento que carece de gran utilidad, revelándose más como un informe de indicaciones generales para alcanzar unos niveles mínimos de calidad en la implementación del programa de Atención prenatal.
	5,3	***Exhaustividad y adecuación de la documentación:*** [Sobre 10]	De forma global la documentación incluida en los anexos es bastante amplia. Sin embargo, se echa en falta el documento de los TdR, así como información relativa al presupuesto (programa y evaluación) y a las entrevistadas desarrolladas (Persona entrevistada, lugar, fecha…). Es necesario indicar también que aunque la forma en la que se estructura el informe a nivel de capítulos es adecuada, la distribución de los epígrafes dentro de los mismos, es manifiestamente mejorable.

0.9 Resumen de fortalezas de la evaluación.

- ⇨ El trabajo de planificación incluyó a una importante diversidad de actores que posibilitó contar con una amplia variedad de perspectivas.
- ⇨ El proceso de evaluación contó con la participación de la población usuaria aportando al trabajo un aspecto cualitativo muy necesario.
- ⇨ Se contemplaron técnicas de análisis cuantitativas y cualitativas.
- ⇨ El proceso de evaluación contó con la participación de gestores del programa del ámbito local y regional que aportaron al trabajo una perspectiva estratégica.
- ⇨ Las conclusiones obtenidas se vinculan con la necesidad de alcanzar unos niveles mínimos de exigencia en la gestión del programa, que garanticen una mejora de la calidad del servicio público evaluado.

0.10 Resumen de debilidades de la evaluación.

Aunque la evaluación adopta particularidades en el campo de la salud y tiene amplia tradición, existen fundamentos teórico-metodológicos que todo proceso evaluativo debe respetar.

- ⇨ Existen confusiones conceptuales y debilidades muy relevantes en el planteamiento y desarrollo de esta evaluación que merman su calidad y limitan la sistematicidad del enjuiciamiento que debe caracterizar toda evaluación.
  - Se plantea un modelo de impactos para evaluar la calidad, cuando son criterios muy diferentes. Es decir, establece previamente el uso de una metodología de evaluación vinculada a la medición de resultados e impactos cuando el objetivo del trabajo debía ser evaluar la calidad de un servicio público. Esto puede ser un error de concepto o una confusión incluso con el propio término.
  - No se explicitan los criterios de valor bajo los que se enjuiciará el programa y se determinará la calidad del servicio.
  - No se establecen las variables o componentes en los que se conceptualiza que se desagrega la calidad y que por tanto, debieran guiar la recolección de información y los indicadores.
  - No se explicita en qué consiste el modelo metodológico que se aplica.
  - No se formulan conclusiones en sentido estricto, sino que se exponen los hallazgos del proceso de manera resumida, si bien el juicio evaluativo que debe caracterizar a las conclusiones, cuando existe, queda diluido en el texto.
  - No se establece una estrategia clara de análisis de la información: triangulación, comparación…

Todos estos motivos hacen complejo aplicarle la lista de criterios y preguntas de esta metaevaluación.

Otras debilidades son:

- ⇨ El modelo de cambio que se desarrolla (aunque presentado gráficamente y con poca explicación, parece razonable y un elemento útil), no se utiliza como eje para articular el proceso evaluativo, es decir, no se usa en la evaluación sino que supone un subproducto de la misma.
- ⇨ Si bien inicialmente se contó con la participación de una amplia variedad de actores de carácter externo al programa, conforme avanzó el proceso de evaluación todas las fuentes de información utilizadas fueron internas (gestores del servicio, expedientes de atención a usuarias, beneficiarias del programa).
- ⇨ El proceso de evaluación se desarrolló por parte de un equipo evaluador que pertenece al mismo organismo que implementa el programa, lo que garantiza el conocimiento experto pero puede generar mermas en su credibilidad.
- ⇨ Los resultados obtenidos no incluyen un análisis de la calidad del servicio basado en la comparativa con información externa al programa. Asimismo, el Informe final de evaluación no contempla un análisis de la eficiencia y eficacia del servicio prestado.
- ⇨ La metodología de evaluación de impacto diseñada no se ha desarrollado de forma correcta. En relación a esto existen importantes carencias relativas al uso de herramientas fundamentales en el desarrollo de una evaluación de impacto (estimación de contra-factual, establecimiento de hipótesis, triangulación de la información recogida,…).

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica	0	Se trata de una evaluación mixta en tanto que es el propio Ministerio de Salud quien implementa y evalúa el programa, con la asistencia para la evaluación de la Caja Costarricense de Seguro Social (CCSS) para la recogida de información. Asimismo, se contó con la asesoría técnica de la Agencia de Cooperación Alemana (GIZ), el Programa de Fomento de las Capacidades en Evaluación (FOCEVAL) y el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN).
		Evaluación interna (100% trabaja en instituciones evaluadas)	1	
		Evaluación mixta (externos y trabajan en instituciones evaluadas)	**5**	
		Evaluación externa (100% externo a instituciones evaluadas)	10	
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y análisis de datos y presentación de resultados	Ninguna / No aplica	0	Las fuentes de información elegidas por el equipo evaluador incluyeron a mujeres usuarias del servicio de Atención Prenatal en el Primer Nivel de Atención (beneficiarias directas), a responsables del nivel regional y local del Ministerio de Salud (gestores del programa) y a médicos de los Equipo Básico de Atención Integral en Salud (EBAIS). (Página 31-32) Asimismo se incluyeron a actores institucionales pertenecientes al Ministerio de Salud, la Caja Costarricense de Seguro Social (CCSS), la Agencia de Cooperación Alemana (GIZ), el Programa de Fomento de las Capacidades en Evaluación (FOCEVAL) y el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN). (Anexo 10 Páginas 139-140).
		Las fuentes elegidas no permiten capturar una pluralidad de perspectivas	1	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas	**5**	
		Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos).	10	
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	***Calificación (numérico)***	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la implementación del proceso de evaluación? ¿Qué grupos? **Fases**: planificación e implementación.	Ninguna / No aplica / No hay información en los documentos consultados	0	La formulación del diseño de evaluación se dio inicialmente en el marco de un acuerdo de colaboración entre Ministerio de Salud, MIDEPLAN y FOCEVAL, para posteriormente integrar a la CCSS. (Página 22). Además, una vez elaborado el borrador del diseño de evaluación y de forma previa al lanzamiento del trabajo de campo, se desarrolló un proceso de sensibilización y capacitación con actores sociales que incluyó dos talleres de capacitación con los equipos de trabajo. (Página 26) Sin embargo, atendiendo a la información presentada en el documento final no se contó en ninguna de esas fases del trabajo con la población beneficiaria del programa.
		No hay participación	1	
		Participan escasa o desigualmente en planificación o implementación	5	
		Participan activamente en la planificación y en la implementación	**10**	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	Dentro del punto 3.6.1.-Preguntas de la evaluación, se explicitan de forma clara las diferentes preguntas de evaluación. Se presentan tres interrogantes principales y cinco interrogantes secundarias vinculadas al contexto de la evaluación, para dar respuesta a los objetivos establecidos. (Páginas 25-26)
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	5	
		Las preguntas son claras y adaptadas	**10**	
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica**	**0**	No se dispone de los TdR.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	Tras la elaboración del borrador del diseño de evaluación y de forma previa al lanzamiento del trabajo de campo, se desarrolló un proceso de sensibilización y capacitación con actores sociales que incluyó dos talleres de capacitación con los equipos de trabajo. (Página 26)
		No hubo un taller de inicio con stakeholders para facilitar su participación	1	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	
		Hubo un taller de inicio con stakeholders	**10**	
2.5		Ninguna / No aplica	0	El informe final de evaluación debería estar disponible a través del siguiente link:

	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	**No son accesibles para el público (ni Informe ni resumen)** Son accesibles para el público (Resumen o extracto) Son accesibles para el público (Informe y resumen)	**1** 5 10	http://ministeriodesalud.go.cr/index.php/material-publicado/descargas/investigaciones/calidad-de-atencion-en-el-primer-nivel-de-atencion?format=html Sin embargo en el momento de realizar este informe el enlace no se encontraba disponible.
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**			
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica No se describe el objetivo de la evaluación (Ni TdR ni Informe) Se describe el objetivo de la evaluación de manera imprecisa **Se describe de manera precisa el objetivo**	0 1 5 **10**	En el punto 3.6 Objetivos, se describe de forma concreta tanto el objetivo general de la evaluación, como su articulación a partir de tres objetivos específicos. (Páginas 24-25)
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica No se describe el objeto de evaluación Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes **Se describe con precisión (al menos los ítems contemplados)**	0 1 5 **10**	Dentro del punto 3 Diseño de la evaluación, se efectúa una descripción precisa del objeto de evaluación. Se detalla información relativa a: • Caracterización de las áreas de salud del estudio. • Problema y objeto de la evaluación. • Cadena de resultados que incluye: Insumos, actividades y productos/servicios. Asimismo, en el punto 2 se detallan el marco conceptual y el jurídico que definen el objeto de la evaluación. Por último, se incluye la Norma Oficial Atención Integral a la Mujer durante el Embarazo de Bajo Riesgo Obstétrico, que es la norma reguladora del programa (Anexo 6). Sin embargo, a lo largo de todo el Informe final se carece de información relativa tanto al monto económico de la evaluación como al volumen financiero del programa. Tan sólo se indican cantidades expresadas en colones, en la norma que se incluye en el Anexo 6.
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica **No se realizó un análisis de stakeholders** Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.) Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	0 **1** 5 10	El diseño metodológico de esta evaluación prestó más atención a la caracterización de las áreas de salud evaluadas que a la realización de un estudio previo de los stakeholders. La población participante en la evaluación está caracterizada en base a variables tales como el haber participado en el programa durante el ejercicio 2012, pero su selección para participar no obedece a ningún estudio ni análisis previo, al menos a partir de la información que se aporta en el informe. Asimismo, en relación a la selección de las áreas evaluadas en el punto 3.2 Caracterización de las Áreas de Salud del estudio, hay que indicar que, si bien inicialmente se presenta un análisis objetivo/justificativo de la selección efectuada, el proceso se vuelve parcial finalmente en tanto que fue la ministra de salud quien determinó finalmente las áreas objeto de estudio. De forma explícita, el texto del Informe final recoge la siguiente afirmación "Finalmente, se consideró el criterio de la señora Ministra de Salud, quien seleccionó dichas áreas de salud."(Páginas 18-19)
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica No se elaboró **Se elaboró pero no se usa en la evaluación o no es concluyente** Se elaboró, se usa en la evaluación y es concluyente	0 1 **5** 10	Uno de los aspectos prioritarios para la implementación de la evaluación fue el diseño de la Cadena de Resultados que define el programa. Este proceso fue realizado por un equipo multidisciplinar donde participaron técnicos del Ministerio de Salud, MIDEPLAN, FOCEVAL y la Caja Costarricense de Seguro Social. (Páginas 22-23) No obstante, no se aprecia su uso en la evaluación, quedando relegada a un subproducto del proceso.
	¿Se formularon hipótesis de impacto concluyentes?	**Ninguna / No aplica** No se formularon Se formularon pero no son concluyentes Se formularon y son concluyentes	**0** 1 5 10	Si bien en la descripción del diseño metodológico se presenta el establecimiento de hipótesis como una herramienta necesaria dentro del análisis cuantitativo, no se recoge dentro del Informe final ninguna hipótesis formulada, de impacto o de proceso, relativa al programa evaluado. Esta situación es coherente en tanto que se trata de una evaluación de la calidad de un servicio público y no del impacto del mismo. Sin embargo, si se atiende a la reflexión presentada en la página 18 del informe "…hay evidencia de que cuidados prenatales adecuados están asociados con al menos cuatro resultados positivos importantes…" habría sido interesante establecer hipótesis de impacto previas y concluyentes, relativas a esos cuatro resultados positivos en la población beneficiaria.
		Ninguna / No aplica	0	

	¿Se diferencia entre los diferentes niveles de impacto? Niveles de impacto:	No se diferencia entre niveles de impacto	1	En el Informe final analizado no se diferencia en ningún momento entre distintos niveles de impacto. Tal y como se indica en la observación anterior, esto puede ser coherente en tanto que se trata de una evaluación de la calidad de un servicio público. Sin embargo, si se atiende a la mencionada reflexión presentada en la página 18 del informe "…hay evidencia de que cuidados prenatales adecuados están asociados con al menos cuatro resultados positivos importantes…" y a qué tal y como se indica en la página 10 "Para el desarrollo de esta evaluación se contó con el Modelo Conceptual y Estratégico de Evaluación de Impacto de Acciones en Salud diseñado por la Unidad de Evaluación de Impacto de Acciones en Salud (UEIAS) del Ministerio de Salud, el cual está alineado con el Manual Gerencial para el Diseño y Ejecución de Evaluaciones Estratégicas de Gobierno de MIDEPLAN" habría sido interesante establecer diferentes niveles de impacto para analizar la calidad del servicio.
		Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)	5	
		Se diferencia con claridad los diferentes niveles de impacto	10	
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica	0	Se describe el contexto de la evaluación aunque no se aprecia su vínculo con la teoría de la intervención del proyecto. A modo de posicionamiento estratégico inicial, se establece en la página 9 la siguiente reflexión "al Ministerio de Salud como ente rector de la Producción Social de la Salud, le compete desarrollar la función estratégica de Evaluación del Impacto de las Acciones en Salud", a partir de la cual se desarrolla la teoría de la intervención.
		No se describe el contexto de la evaluación	1	
		Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención	**5**	
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	10	
3.2	**Diseño y métodos**			
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (o los TdR no explicitan los intereses de la evaluación)	0	No se dispone de los TdR para valorar los intereses de la evaluación, y al ser mixta es probable que no existan. En el Informe se anuncia que se aplica Modelo Conceptual y Estratégico de Evaluación de Impacto de Acciones en Salud diseñado por la Unidad de Evaluación de Impacto de Acciones en Salud (UEIAS) del Ministerio de Salud. No obstante, no se desarrolla ni se aprecia los fundamentos o contenidos de dicho modelo, ni se aborda la medición del impacto en sí. La evaluación se centra evaluar la calidad del servicio de atención prenatal, indicándose en la página 9 que: *"La calidad de la atención, se refiere a un conjunto de múltiples interrelaciones en aspectos como: equidad, eficiencia, eficacia, la oportunidad de la atención; así como la humanización de los servicios y la satisfacción de los usuarios y proveedores."* No obstante no es fácil aprecia el diseño metodológico en sí, pues el Informe explica detalladamente las técnicas de investigación que aplica para dar respuestas a preguntas relacionadas con el cumplimiento de una serie de normas y protocolos de salud, más que en el enjuiciamiento de la bondad y el mérito de una política pública. A este respecto puede decirse que las técnicas que se aplican se adecúan para ofrecerles una respuesta, pero no establecen cómo juzgarán la calidad, en qué componentes o categorías de análisis se desagrega esta, o en qué medida y cómo se asocia a la eficacia o a la eficiencia como se anuncia. Por tanto se concluye que el diseño permite valorar el cumplimiento normativo y de protocolos pero presenta debilidades en cuanto al establecimiento de criterios de valor explícitos sobre los que se asiente la emisión sistemática del juicio evaluativo.
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	**5**	
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	10	
3.2.2	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	Ninguna / No aplica	0	Dentro del punto 3 Diseño de la evaluación, se explicitan los métodos de evaluación aplicados (cualitativos y cuantitativos), se caracterizan las fuentes de información utilizadas, y se indican los instrumentos de análisis usados. Sin embargo, esto no quiere decir que sean los más adecuados en cantidad y calidad para desarrollar de la mejor manera esta evaluación.
		No se explicitan ni justifican el diseño y los métodos aplicados	1	
		Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados	**5**	
		Se explicitan y justifican suficientemente el diseño y los métodos aplicados	10	
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	**Ninguna / No aplica**	**0**	En base a la Información presentada en el documento final se puede afirmar que la situación contrafactual no fue un aspecto considerado por el equipo evaluador. En ningún momento se hace mención dentro del Informe final a la construcción de grupos de comparación.
		No se consideró lo contrafactual	**1**	
		Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró	5	
		Se consideró lo contrafactual y se especificó cómo	10	
3.2.4		Ninguna / No aplica	0	Dentro del punto 3 Diseño de la evaluación, se presentan métodos de análisis que cubren ambos aspectos:

	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Los métodos aplicados no cubren aspectos cuanti ni cuali	1	3.7.1-Enfoque cuantitativo (Página 27) y 3.7.2 Enfoque cualitativo (Página 32).
		Los métodos aplicados cubren aspectos cuanti o cuali	5	
		Los métodos aplicados cubren aspectos cuanti y cuali	**10**	
3.2.5	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) Entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica	0	
		Se aplicó sólo un método	1	
		Se aplicaron dos métodos	**5**	
		Se aplicaron una pluralidad de métodos (más de dos)	10	Técnica / Método: Entrevistas – Tipo: guiadas – Nº Ediciones: 17 – Nº Participantes: 17 Grupos – focales – 4 – 32 Encuestas – Face to face, telefónicas, online Se desarrollaron de forma previa al lanzamiento de la evaluación 2 talleres de capacitación con los equipos de trabajo (Página 26) Para el trabajo cuantitativo, antes de proceder a cumplimentar los Instrumentos de recogida de información cuantitativa 1 y 2 (Anexos 1 y 2) se realizaron 17 entrevistas a médicos directores de los EBAIS. (Página 28) Para el análisis cualitativo se realizaron 4 Grupos focales en los que participaron un total de 32mujeres pertenecientes a 4 EBAIS (6+6+6+14) (Páginas 33 y 35)
3.3	**Procedimiento de selección**			
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica	0	Dentro del epígrafe 3.7.1 Enfoque cualitativo, se determina tanto el procedimiento para cumplimentar los instrumentos de recogida de formación cuantitativa 1 y 2, como la determinación de la población meta a partir de la que extraer la muestra objeto de estudio. (Páginas 27-31) (ver Tablas Nº3 y Nº4) Dentro del epígrafe 3.7.2 Enfoque cualitativo, se especifican los criterios de selección de la población participante en los Grupos focales, así como los responsables de efectuar dicha selección. (Página 32-33)
		No se explicitan los procedimientos ni criterios de selección	1	
		Se explicitan en algunos casos (técnicas)	5	
		Se explicitan los procedimientos y criterios de selección	**10**	
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? Se entiende que aplica esta pregunta para las Encuestas	Ninguna / No aplica	0	La Tabla Nº3 contiene la Distribución de la muestra de acuerdo con la Consulta Prenatal, según Áreas de Salud. Se trata de una muestra de 781 consultas a mujeres beneficiarias del programa con las siguientes características: • Nivel de confianza en las estimaciones del 95,0%. • Tasa máxima de consulta del 50,0%. • Margen de no respuesta del 3,0%. • Margen de error máximo permisible de ± 3,5%. (Páginas 30-31).
		La muestra no es adecuada	1	
		La muestra no es adecuada o el error no es aceptable	5	
		La muestra sí es adecuada y el error aceptable	**10**	
3.4	**Análisis de datos**			
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica	0	El análisis de la información se hizo de forma aislada, separando los aspectos cuantitativos de los cualitativos. Tanto el punto 4 Análisis de la información, como el punto 5 relativo a Conclusiones presentan la información separadamente (cuantitativo-cualitativo). De hecho, se presenta los resultados obtenidos con los instrumentos 1 y 2 de forma aislada (cuantitativo) y del mismo modo el análisis de los 4 grupos focales (cualitativo). Tan sólo en un par de conclusiones de carácter cuantitativo (3y 6) se matizó el resultado obtenido con información recabada en los Grupos focales.
		No se triangulan datos cuanti-cuali	**1**	
		Se triangulan datos cuanti-cuali pero no de manera generalizada	5	
		Se triangulan datos cuanti-cuali de manera generalizada	10	
3.4.2	¿Se cuantifican la respuesta siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados…?	Ninguna / No aplica	0	Tanto en el punto 4 Análisis de la información, como en el punto 5 Conclusiones, se detallan y ponderan las respuestas obtenidas a partir del uso de los instrumentos de análisis cuantitativo: 1 Planificación y Organización de la Atención Prenatal y 2 Revisión de Expedientes de Atención Prenatal.
		No se cuantifican las respuestas cuando es posible	1	
		Se cuantifican las respuestas a veces cuando es posible	5	
		Se cuantifican las respuestas de manera generalizada cuando es posible	**10**	
3.4.3		Ninguna / No aplica	0	

	¿Se usaron datos del monitoreo del proyecto/programa?	No se usaron datos del monitoreo	1	La fuente que se utilizó principalmente para obtener la información cuantitativa necesaria para desarrollar esta evaluación fue el registro de expedientes personales de las usuarias que asistieron a la atención prenatal durante el año 2012. (Página 28) Se entiende que este registro de expedientes personales de las mujeres que acuden a consulta, es parte del sistema de monitoreo del programa de Atención prenatal en el primer nivel de atención.
		Se usaron datos del monitoreo de manera insuficiente	5	
		Se usaron datos del monitoreo de manera suficiente	**10**	
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica	0	El equipo evaluador no recurrió a fuentes de datos externos para efectuar el análisis cuantitativo ni cualitativo.
		No se recurrió a fuentes externas	**1**	
		Se recurrió a fuentes externas de manera insuficiente	5	
		Se recurrió a fuentes externas de manera suficiente	10	
3.5	**Criterios de evaluación**			
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			Es esta una evaluación que no establece explícitamente criterios de valor como base para juzgar la bondad, el valor y mérito de esta intervención pública. Desde un punto de vista metodológico existe debate en torno a si es necesario establecerlos o no para evaluar una política pública, de hecho Stake (2007) desarrolla una obra confrontando dos modos de evaluar: evaluación comprensiva y evaluación basada en estándares (o criterios). En cualquier caso, aplicar a esta evaluación estas preguntas plantea una cuestión: responder que no aplican estos criterios a esta evaluación o responder que no los expone de manera concluyente y comprensible (con la consiguiente merma en su calificación). En esta metaevaluación se ha optado por considerar el grado de respuesta que ofrece a cada uno de los criterios que se estimen le son aplicables (a excepción del impacto y la sostenibilidad). Los criterios de eficacia y eficiencia son nombrados explícitamente como componentes de la calidad según indica el propio informe en su marco conceptual.
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	Dentro del punto 2 relativo al marco conceptual, el Informe final describe la importancia del proyecto para la región y el país en general, y para la salud de las beneficiarias del servicio de Atención prenatal en particular. Se pone por tanto de manifiesto en el informe final la relevancia y pertinencia de un programa público como el evaluado. No obstante, cabe apuntar que esto se expone como contextualización de la evaluación y como delimitación del objeto que se evalúa, no suponiendo en ningún caso la emisión de un juicio de valor por parte del equipo evaluador a cerca de los criterios de relevancia y pertinencia –que es lo que interpela esta pregunta-, por lo que se considera que no expone esta pertinencia y relevancia de forma concluyente.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	**5**	
		Lo expone de forma concluyente y comprensible	10	
	b) logró las metas definidas	Ninguna / No aplica	0	Al tratarse de una evaluación de la calidad de un servicio público no se contempla el análisis relativo al logro de sus metas u objetivos. No obstante, puede considerarse que las metas son el cumplimiento de los requerimientos normativos y de protocolo que se evalúan.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	
		Lo expone de forma concluyente y comprensible	10	
	c) es eficiente	**Ninguna / No aplica**	**0**	Dentro del documento final no se encuentra información concluyente relativa a un análisis de eficiencia del programa. Se considera que no aplica este criterio
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	d) causó impactos (outcome e impacto) deseados y no deseados.	**Ninguna / No aplica**	**0**	El Informe final no expone de forma concluyente los resultados del programa a nivel de impactos (outcomes). En relación a este ítem hay que indicar que no se trata de una evaluación de resultados o impactos.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	Tal y como se indica en el apartado anterior, al tratarse de una evaluación de la calidad del servicio, no se expone de forma concluyente en el Informe final los resultados del programa a nivel de impactos (outcomes).
		No lo expone	1	

		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	La fuente que se utilizó principalmente para obtener la información cuantitativa necesaria para desarrollar esta evaluación fue el registro de expedientes personales de las usuarias que asistieron a la atención prenatal durante el año 2012. (Página 28) Se entiende que este registro de expedientes personales de las mujeres que acuden a consulta, es parte del sistema de monitoreo del programa. Sin embargo, en el Informe no se detalla información relativa a la orientación para la evaluación de este sistema de monitoreo. Se trataría por tanto de un sistema enfocado a la gestión, utilizado en este caso para el análisis de la calidad del servicio.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	**10**	
3.5.2	g) ¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	Desde un punto de vista cuantitativo, los resultados presentados en el Informe final no van más allá de un análisis descriptivo de la situación actual de los EBAIS y de una síntesis del seguimiento efectuado a una muestra representativa de las mujeres que en 2012 fueron beneficiarias del servicio de Atención prenatal. Por otra parte, focalizando el análisis a través de una óptica cualitativa, se detallan las valoraciones personales y subjetivas de 32 mujeres participantes en 4 Grupos focales, sobre el servicio recibido como beneficiarias del programa de Atención prenatal. Si bien se trata de unas herramientas de gran utilidad en el desarrollo de un proceso evaluativo, su uso exclusivo, sin complementarlas con otras técnicas de análisis, no permite detectar cambios en la población beneficiaria ni inferir los mismos al proyecto evaluado.
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	**1**	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	5	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	No se dispone de los TdR. No obstante el Informe establece en su apartado 3.6.1 las preguntas de evaluación. Es compleja la respuesta a este ítem, pues el informe no presenta sus conclusiones articuladas por preguntas de evaluación, ni tampoco por criterios. Tampoco puede decirse que presente conclusiones únicas, fruto del enjuiciamiento que formulan los evaluadores sobre los hallazgos, dado que no existe triangulación y se presentan un listado conclusiones por técnicas cuantitativas y otro por las cualitativas. Esto supone una limitación en su calidad. Por ello, se le otorga una puntuación intermedia en este ítem, por la dificultad que entraña tratar de obtener una respuesta global a las preguntas evaluativas, no tanto porque no pueda ofrecerse, aunque en todo caso, esto exige al lector un esfuerzo adicional y diluye el propio enjuiciamiento del equipo evaluador.
		No responde adecuadamente a ninguna pregunta	1	
		No responde adecuadamente a alguna/s pregunta/s	**5**	
		Responde adecuadamente a todas las preguntas	10	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	Los resultados (cuantitativos y cualitativos) presentados en el punto 5 relativo a Conclusiones no van más allá de un análisis descriptivo del programa que incluye valoraciones del estado de los EBAIS y de la calidad del servicio según una muestra de la población beneficiaria.
		No se describen puntos fuertes y débiles	**1**	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	10	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	La información se presenta de forma descriptiva en el caso de los resultados obtenidos (Conclusiones) y propositiva en relación a las recomendaciones generales efectuadas.
		No se distinguen con claridad declaraciones descriptivas y valorativas	**1**	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	10	
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	La fuente de información cualitativa (mujeres participantes en los Grupos focales) es primaria, y atendiendo al proceso de selección de las participantes es confiable para el fin que se utilizan. Las fuentes de información cuantitativa son gestores del programa y registros públicos de carácter primario y confiable: Médicos directores de los EBAIS que participan en la evaluación; y Expedientes médicos de mujeres que recibieron atención prenatal.
		No se especifica	1	
		Se especifica en ocasiones	5	
		Se especifica en general	**10**	
4.5		Ninguna / No aplica	0	Las conclusiones se vinculan a los resultados obtenidos.

	¿Las conclusiones están basadas en los resultados?	**Las conclusiones no están basadas en los resultados**	**1**	Sin embargo, hay que indicar que estas conclusiones no se establecen a resultas de un proceso de triangulación de la información cuantitativa y cualitativa obtenida, sino que representan un resumen numerado de los principales resultados revelados en el proceso de evaluación.
		Alguna o algunas conclusiones no están basadas en resultados	5	
		Todas las conclusiones están basadas en los resultados	10	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	El Informe final presenta un número mínimo de recomendaciones si se atiende al elevado número de conclusiones registradas (7 frente a 21). Asimismo, las recomendaciones presentadas son excesivamente genéricas, focalizadas en mejorar los aspectos negativos de la gestión del programa y la falta de recursos, registrados durante el trabajo evaluativo.
		No se llega ni de forma lógica ni comprensible	**1**	
		Se llega de forma lógica o comprensible pero no las dos	5	
		Se llega de forma lógica y comprensible	10	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	El listado de recomendaciones propuesto no es relevante ni específico. Asimismo, las propuestas de mejora presentadas son excesivamente genéricas, indicando aspectos tan generales para alcanzar un nivel mínimo de calidad en la implementación de un programa público como: • La necesidad de dotar de equipos y materiales a los EBAIS. • La obligatoriedad de cumplir determinados procedimientos y normativa, que regulan el programa.
		No son relevantes ni específicas	**1**	
		Son relevantes o específicas pero no las dos	5	
		Son relevantes y específicas	8-9-10	
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	De forma introductoria, el informe incluye un resumen ejecutivo bien estructurado, claramente legible y comprensible, en el que se presentan de forma sintética los aspectos esenciales de la evaluación. (Páginas 9-10)
		El Informe no contiene un resumen	1	
		El informe contiene un resumen estructurado	5	
		El informe contiene un resumen estructurado, bien legible/comprensible	**10**	
5.	**Exhaustividad y adecuación de la documentación**			
5.1	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Ninguna / No aplica	0	No se dispone de los TdR.
		Los TdR no son accesibles en anexos	**1**	
		Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	
		Los TdR son accesibles en la lengua común de trabajo o en una traducción	10	
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	Se incluyen actores institucionales participantes (Anexo 10) y cronograma de trabajo (Anexo 12). Aunque se incluyen los documentos para la recogida de información cuantitativa (Anexo 1 y Anexo 2) y una guía para el desarrollo de los talleres focales (Anexo 3), no se detallan ni las personas entrevistadas ni los participantes en los talleres focales.
		No se incluyen en anexos	1	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	**5**	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	En los anexos se incluyen los documentos para la recogida de información cuantitativa (Anexo 1 y Anexo 2) y una guía para el desarrollo de los talleres focales (Anexo 3). Asimismo, se presenta el detalle del material de consulta específico relacionado con el programa utilizado en la evaluación (Anexos 4 y 5), la Norma Oficial Atención Integral a la Mujer durante el Embarazo de Bajo Riesgo Obstétrico, que es reguladora del programa (Anexo 6), y otras normas costarricenses focalizadas en la salud de la madre y el niño/a (Anexo 7).
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	**10**	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve	Ninguna / No aplica	0	El informe final de evaluación presenta una estructura correcta para facilitar su lectura, pero mejorable. Si bien la forma en la que se estructura el informe a nivel de capítulos es adecuada, la distribución de los epígrafes dentro de
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	

	facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	**El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura o elementos facilitadores).**	**5**	los mismos, es manifiestamente mejorable. En este sentido hay que indicar que no se diferencia de manera clara entre los diferentes niveles de los epígrafes que componen un capítulo, no facilitando este hecho su lectura. No se incluyen resúmenes de los capítulos o epígrafes introductorios. Sí se presentan gráficos y tablas en un número suficiente.
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	10	

4.4.11 Evaluación 011: PyA-Cáncer

0.1 Nombre del proyecto / programa evaluado	**EVALUACIÓN PARTICIPATIVA DE LOS SERVICIOS DE PREVENCIÓN Y ATENCIÓN DEL CÁNCER EN VALLE LA ESTRELLA, PROVINCIA DE LIMÓN, COSTA RICA FEBRERO DE 2017.**	
0.2 Institución que implementa	Caja Costarricense del Seguro Social	
0.3 Evaluadores	*Equipo de evaluación* Caja Costarricense de Seguro Social (Karol Aguilar, Virginia Venegas) Juntas de Salud: - Matina (Esther Barquero); - Siquirres (Norma Barr); - Guácimo (Fressy Calderón); - Valle de la Estrella (Herman Fernánder; Susana Olivares; Olga Ramírez); - Limón (Eugenia Romero)	*Equipo de apoyo* - Defensoría de los Habitantes (Hannia Silesky; Rodolfo Vargas) - MIDEPLAN (Ericka Valerio) *Expertos consultores independientes* - Experto evaluación participativa (Esteban Tapella) - Facilitadora sesiones de trabajo (Karla Salazar) *Coordinación* - DEval (Andrea Meneses; Juan Carlos Sanz)
0.4 Monto del proyecto/ programa	No especificado en el Informe final de evaluación.	
0.5 Monto para la evaluación	No especificado en el Informe final de evaluación.	

0.6 Descripción del proyecto/programa evaluado (un párrafo)

Los servicios de prevención y atención del cáncer que presta la Caja en Valle de la Estrella encuentran su anclaje estratégico en el Plan Nacional para la Prevención y Control del Cáncer 2012 – 2017 (Ministerio de Salud, 2012), cuyo propósito es lograr el control del cáncer mediante un enfoque sistémico, un abordaje integral, una gestión coordinada y la participación de los actores sociales. En este sentido destaca el rol auxiliar de las Juntas de Salud, a través de las que se canaliza dicha participación.

La prestación de estos servicios se articula en dos ámbitos: prevención y atención en tres niveles (primaria, secundaria y terciaria).

Los **servicios de prevención** se prestan en los centros de salud (establecimientos de atención ambulatoria, Equipos Básicos de Atención Integral en Salud (EBAIS) y consultorios de visita periódica). Se cuenta con el apoyo de enlaces de oncología, Asistentes Técnicos de Atención Sanitaria Primaria (ATAP) y otros actores (estudiantes de medicina o trabajadores sociales). Se realizan campañas en distintos medios de difusión, charlas informativas y, ocasionalmente, ferias de la salud.

La **atención primaria** corresponde también a los centros de salud, que ofrecen servicios de consulta externa y detección precoz de enfermedades (entre ellas el cáncer), curaciones y rehabilitaciones de menor complejidad. El personal médico aplica un formulario sobre factores de riesgo y ofrece información; derivando al nivel secundario posibles casos de cáncer y apoya a pacientes y familias.

La **atención secundaria** corresponde a 13 hospitales periféricos y 7 regionales (Tony Facio, en Limón, para los pacientes de Valle de la Estrella). Prestan apoyo al primer nivel, coordinan de enlaces de oncología; ofrecen intervenciones ambulatorias y hospitalarias para especialidades básicas y realizan cirugía poco compleja y pruebas que de ser positivas se derivan al tercer nivel.

La **atención terciaria** corresponde a 3 hospitales nacionales generales y 6 especializados, siendo el Hospital Calderón Guardia, en San José el que atiende a la población de Valle de la Estrella. Presta servicios ambulatorios y de hospitalización más complejos o que requieren alta tecnología y especialización.

0.7 Descripción de la modalidad de implementación/apoyo de FOCEVAL (máximo media página)

Evaluación participativa impulsada por el Consejo Regional de Juntas de Salud y realizada por un equipo de evaluación con representación de las Juntas de Salud y de la Caja Costarricense del Seguro Social que contó con el apoyo de MIDEPLAN y la Defensoría de Habitantes, la asistencia técnica de consultores en evalaución participativa y facilitación de procesos participativos, y la coordinación de FOCEVAL-DEval.

A partir de la propuesta de la propuesta del Grupo de Coordinación de FOCEVAL de vincular evaluación y participación ciudadana (junio de 2014) se inicia una línea de actividad para promover evalauciones participativas. Esta línea se inicia con un Encuentro con organizaciones sociales para conversar sobre evaluación (julio de 2015). Continúa con una capacitación en evaluación participativa dirigida a estas organizaciones (noviembre de 2015). Y, deriva, en enero de 2016 en la convocatoria de evalauciones participativas dirigidas a estas organizaciones sociales ofreciéndoles apoyo para elaborar sus propuestas. En esta convocatoria fue seleccionada la propuesta presentada por Consejo Regional de Juntas de Salud de la Región Huétar Atlántico, dando inicio en junio de 2016.

Esta modalidad se concreta en apoyar iniciativas en las que los propios stakeholders demandan la evaluación y conforman el equipo evaluador, contando con el apoyo de MIDEPLAN, ofreciendo FOCEVAL asesorías especializadas y asumiendo DEval la coordinación. En esta evaluación el equipo evaluador (Junta de Salud y CCSS) recibió asesoramiento experto (en evaluación participativa y en facilitación de procesos participativos), contó con el apoyo de MIDEPLAN y la Defensoría de Habitantes, así como con la coordinación de FOCEVAL. La modalidad de apoyo estuvo focalizada en desarrollar las capacidades de las Juntas de Salud y de la CCSS para facilitar su apropiación del proceso, garantizando la sistematicidad y el rigor de la evaluación a paritr de los apoyos, asesorías y la coordinación del proceso.

0.8 Cualificación cuantitativa (con comentarios) según criterios			
73,1 [Sobre 100]	**5**	***Imparcialidad e independencia:*** [Sobre 10]	La evaluación obtiene una puntuación intermedia en este criterio, al optar por un equipo mixto y "… y recurrir a un número limitado de fuentes que además tuvieron que ser adaptadas para que el equipo evaluador pudiera apropiarse de las mismas"..
	10	***Equidad, participación y transparencia:*** [Sobre 10]	Esta evaluación obtiene la mayor puntuación posible en términos de equidad y participación. Se apoya para ello en una implicación sostenida de los stakeholders a lo largo del proceso; se formulan preguntas claras y adaptas precisamente desde la plataforma que constituye el equipo evaluador donde dichos stakeholders están representados; y no sólo hubo talleres de inicio, sino que el propio lanzamiento de la evaluación estuvo precedido de un proceso de desarrollo de capacidades, que se mantuvo a lo largo de la evaluación en las sesiones/talleres de trabajo del equipo evaluador. También en términos de transparencia, la aceesibilidad a través de la web del informe y su resumen ejecutivo supone un factor positivo.
	31	***Credibilidad y exactitud:*** [Sobre 50]	Tomando los subcriterios que vertebran el análisis de credibilidad y exactitud, esta evaluación obtiene buenos registros en cuanto a la delimitación de Metas, objetivos y modelos de cambio (8 sobre 10), también en su Diseño y métodos (8,8 sobre 10) y en la explicitación de sus Procedimientos de selección (10 sobre 10), pues justifica sobradamente su diseño así como las técnicas que aplica y los criterios de selección desplegados en ellas. Estos buenos registros contrastan con sus logros en cuanto a Análisis de datos (4 sobre 10), donde le penaliza el hecho de que sea exclusivamente cualitativa no triangule datos cuantitativos con ellos, no contemple datos del monitoreo y hace un uso muy limitado de fuentes externas. Asimismo, sus registros en cuanto al tratamiento de los Criterios de evaluación (5,2 sobre 10) está derivado de que si bien responde a todas las preguntas de evaluación a partir de los hallazgos, el juicio evaluativo queda diluido en una suerte de síntesis de hallazgos.
	18,8	***Calidad del informe y utilidad:*** [Sobre 20]	El informe obtiene una muy valoración sobre su calidad y utilidad, al dar respuesta suficiente a las preguntas de evaluación, presentando de manera balanceada puntos fuertes y débiles, y existiendo una trazabilidad adecuada entre hallazgos, conclusiones y recomendaciones.
	8,3	***Exhaustividad y adecuación de la documentación:*** [Sobre 10]	En cuanto a la documenta que aporta en sus anexos, contempla la matriz de evaluación, las guias de los instrumentos aplicados y las fuentes consultadas, así como los principales datos del cronograma del trabajo de campo (que también aparecen a lo largo del informe), aunque no se detallan las personas participantes o las organizaciones a las que pertencen en todos los casos, siendo esto especialmente necesario en las entrevistas para poder valorar su cobertura de perfiles específicos.

0.9 Resumen de fortalezas de la evaluación.

- El desarrollo de capacidades en evaluación en el seno del equipo evaluador conformado por las Juntas de Salud y la CCSS.
- La utilidad potencial del informe para las Juntas de Salud y la CCSS, en cuanto a su capacidad para inducir mejoras y la factibilidad de aplicar las recomendaciones por el hecho de apreciarse un importante nivel de apropiación del proceso.
- La sistematicidad y el rigor del planteamiento evaluativo, en el que no se aprecia que el conocimiento experto y directo del equipo evalaudor interno (Juntas de Salud y CCSS) haya sido un componente clave para fundamentar conclusiones no basadas en los hallazgos del proceso, lo cual siempre constituye un riesgo cuando se hace una evaluación interna (pre-juicio).
- La prudencia de las preguntas de evaluación, en las que se aprecia que ante el reconocimiento de las limitaciones del diseño evaluativo, se ha desplegado un proceso dinámico de focalización y priorización, ajustando expectativas.

0.10 Resumen de debilidades de la evaluación.

- Los métodos y técnicas cualitativas encuentran su representatividad cuando alcanzan la saturación de los discursos recabados, sin embargo en este informe de evaluación no se aporta información suficiente para juzgar este aspecto clave.
  - Esto es así especialmente en el caso de las entrevistas y los grupos focales, donde tampoco se puede apreciar en qué medida abarcan cada uno de los perfiles relevantes en los diferentes niveles y ámbitos de la prevención y la atención del cáncer, tanto en la vertiente de prestación directa del servicio en lo local como en cuanto al componente institucional desde la perspectiva de la CCSS.
  - Esto también aplica al juego sobre mitos del cáncer, que se aprecia más como subproducto del proceso (útil a futuro) que como soporte de hallazgos que faciliten juicios concluyentes sobre grado de conocimiento, especialmente por su limitada aplicación.
- El componente valorativo de la evaluación, centrando en la formulación del juicio de valor que debe caracterizar a las conclusiones, ha quedado diluido en este informe, en el que suele recurrirse a un balance-síntesis de las percepciones recabadas, especialmente en aquellos criterios en los que se aprecia una menor vertebración a partir de sus principales dimensiones (eficacia y eficiencia frente a pertinencia).
- El escaso uso de fuentes externas e información del monitoreo que aparece condicionado por la premisa de la evaluación participativa sobre que sea el propio equipo evaluador (Juntas de Salud y CCSS) el que procese la información y tome las decisiones, lo que exige acotar, derivadando en la renuncia a información muy relevante para la solvencia de los hallazgos.
- La orientación de la evaluación hacia el desarrollo de capacidades de un equipo evaluador interno que conoce profundamente el contexto y los servicios evaluados, ha derivado en una menor precoupación en el informe final por ofrecer información relevante para juzgar la solvencia de los hallazgos.

1.	Imparcialidad e Independencia	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
1.1	¿Se trata de una evaluación externa o interna (miembros del equipo evaluador trabajan en las instituciones evaluadas)?	Ninguna / No aplica Evaluación interna (100% trabaja en instituciones evaluadas) **Evaluación mixta (externos y trabajan en instituciones evaluadas)** Evaluación externa (100% externo a instituciones evaluadas)	0 1 **5** 10	Se trata de una evaluación mixta, en tanto que: - La Caja Costarricense, que brinda los servicios que se evalúan, participa en el equipo evaluador (con dos integrantes). Estas integrantes son responbales de la prestación del servicio a nivel local, fungiendo una como Coordinadora Regional de Oncología, y otra como Coordinadora del EBAIS y enlace local de oncología (Pág.28). - Las Juntas de Salud conforman el resto del equipo de evaluación (7 integrantes que representan a los usuarios de los servicios de salud de la zona). Estas Juntas forman parte del entramado institucional dispuesto y regulado normativamente como órganos auxiliares de la Caja para estos fines. Son estas Juntas de Salud además las que se erigen en este caso como promotoras de la evaluación a través de su presentación a la convocatoria de evaluaciones participativas lanzada por FOCEVAL, la Defensoría de los Habitantes y MIDEPLAN. - El equipo de evaluación (CCSS y Juntas de Salud) recibe acompañamiento externo especializado. • El asesoramiento especializado que recibe el equipo evaluador, en evaluación participativa y para facilitar el proceso, es externo bajo la figura de consultores independientes. • La coordinación del proceso de evaluación es también externa y corresponde a DEval-FOCEVAL. • El apoyo a este proceso es brindado por órganos externos : la Defensoría de los Habitantes y MIDEPLAN. Hay que tener en cuenta que el porpio informe ubica esta evaluación entre los "Enfoques Colaborativos de Evaluación, donde los o las evaluadoras trabajan en colaboración con los miembros de la comunidad para realizar la evaluación conjuntamente" (Pág.22).
1.2	¿Se eligieron las fuentes de información de forma que permiten capturar una pluralidad de perspectivas en el levantamiento y el análisis de datos, y en la presentación de los resultados? **Etapas**: levantamiento y análisis de datos y presentación de resultados	Ninguna / No aplica Las fuentes elegidas no permiten capturar una pluralidad de perspectivas **Las fuentes elegidas permiten capturar la pluralidad de perspectivas pero no se aplican en todas las etapas** Las fuentes elegidas permiten capturar la pluralidad de perspectivas en todas las etapas (levantamiento y análisis de datos).	0 1 **5** 10	En cuanto a las fuentes elegidas en el propio Informe se indica lo siguiente: "Junto a la revisión de la documentación que se ha señalado en el capítulo de Referencias de este informe, la principal fuente de información de la evaluación han sido los diferentes grupos de actores vinculados a los servicios de prevención y atención del cáncer en Valle de la Estrella" (Pág.70). "La recolección de los datos se planificó de acuerdo con el tiempo y los recursos disponibles, pero siempre procurando contar con la opinión de la mayor diversidad de actores y puntos de vista posibles" (Pág.26). "Las personas consultadas se pueden clasificar en tres grupos de actores: profesionales de la salud, personas usuarias del sistema de salud y representantes de las Juntas de Salud" (Pág.26). Entre los profesionales de salud consultados, los perfiles concretos que se explicitan en el informe son los Enlaces de oncología y personal médico del Valle de la Estrella (Pág.70). Y entre las personas usuarias, los perfiles que se especifican en el informe son personas diagnosticadas con cáncer y sus familiares, así como otras personas de determinadas comunidades del Valle de la Estrella (Pág.70). No se especifica en el Informe en qué medida los agentes consultados como profesionales del sector salud representan los diferentes servicios de prevención y atención en sus diferentes niveles (primaria, secundaria, terciaria), ni en qué medida integran los diferentes niveles jerárquicos desde los que se articula o contempla en la CCSS la prestación de estos servicios en la zona (representación institucional), o si entre los profesionales contemplan iniciativas y clínicas privadas que también operan en este ámbito y que aparecen referidas en el informe. Estas omisiones, unido a que no hay un análisis explícito de stakholders, dificultan emitir un juicio sobre la pluralidad de perspectivas de la evaluación. En cuanto a las fuentes secundarias (estadísticas oficiales, datos de monitoreo –listas de espera, prevalencias,…), presentan un rol tangencial en la evaluación. Su escaso uso no facilita la contextualización del objeto de evaluación, ni contribuye al análisis de pertinencia, eficacia o eficiencia. En cuanto al rol de las diversas fuentes contempladas, destaca el hecho de que su pluralidad se pone de manifiesto en el informe tanto en el levamiento de datos como en el análisis y en las iniciativas contempladas de presentación de resultados.
2.	**Equidad, participación y transparencia**	**Calificación (texto)**	***Calificación (numérico)***	**Observaciones cualitativas**
2.1	¿Estaban involucrados activamente los grupos meta y otros grupos interesados en la planificación y la	Ninguna / No aplica / No hay información en los documentos consultados No hay participación	0 1	Esta evaluación se caracteriza precisamente por la implicación de los grupos meta y otros grupos a lo largo de todo el proceso de evaluación, siendo su seña identitaria.

	implementación del proceso de evaluación? ¿Qué grupos? **Fases**: planificación e implementación.	Participan escasa o desigualmente en planificación o implementación	5	▪ Grupos implicados en todas las fases: CCSS (responasbles de la prestación del servicio a nivel local) y Juntas de Salud (representantes de grupos meta) como integrantes activos del Equipo de Evaluación ▪ Otros grupos implicados durante la implementación como informantes en diversas técnicas: profesionales médicos, personas usuarias del servicio y de la comunidad, personas diagnosticadas con cáncer o familiares. [Es esta la evaluación que en mayor medida responde a esta pregunta de las que aquí se metaevalúan]
		Participan activamente en la planificación y en la implementación	**10**	
2.2	¿Las preguntas de evaluación fueron formuladas de forma clara y adaptadas al contexto del proyecto?	Ninguna / No aplica	0	Las preguntas de evaluación reciben una atención especial en esta evalaución, tanto sobre su significado, importancia y vínculo con los objetivos específicos, como sobre la explicación del proceso de su formulación y finalmente se expone el detalle de las preguntas concretas que se formularon en la matriz inicial (Anexo 1). Junto a ello, la exposición de los hallazgos se articula sobre los criterios de evalaución y en cada uno de ellos se especifican las preguntas que los concretan, facilitando la trazabilidad del análisis y y su comprensión. Se entiendo (aunque no se explicita) que estas preguntas que se concretan en los hallazgos corresponden a las finalmente priorizadas (pues hay otras en la matriz del anexo que no reciben atención o se atienden en menor medida, poniendo de relieve un proceso dinámico y constructivista en su concepción. Todas las preguntas formuladas, iniciales y finales, son claras y están adaptadas al contexto del servicio que se evalua.
		Las preguntas ni son claras ni adaptadas	1	
		Las preguntas o son claras o adaptadas	5	
		Las preguntas son claras y adaptadas	**10**	
2.3	¿Los TdR fueron formulados de forma clara y comprensible?	**Ninguna / No aplica**	**0**	No se dispone de TdR. Esta evalaución surge a partir de una convocatoria a organizaciones sociales para realizar evalauciones participativas, por lo que no existen TdR en sentido estricto. Cabría concebir esta convocatoria, la propuesta presentada por el Consejo Regional de Juntas de Salud de la Región Huétar Atlántico y los criterios que llevaron a seleccionarla, como los términos que orientan el encargo, junto a la matriz inicial de evalaución. En este sentido, hubiese sido deseable contemplar en el informe alguna referencia, en el texto o en los anexos, a los términos concretos de esta convocatoria o propuesta de forma que se permitiese valorar en qué medida el encargo de evaluación fue claro y comprensible.
		Los TdR presentan una formulación ni clara (estructura) ni comprensible (redacción)	1	
		Los TdR presentan una formulación clara (estructura) o comprensible (redacción)	5	
		Los TdR presentan una formulación clara (estructura) y comprensible (redacción)	10	
2.4	¿Tuvo lugar un taller de inicio? ¿Qué grupos de interés fueron involucrados?	Ninguna / No aplica	0	Esta evaluación es paradigmática en este sentido, pues no sólo se contemplaron talleres de inicio como los que refiere esta pregunta para implicar a los stakeholders cuando va a dar comienzo la evaluación sino que contempló incluso un proceso previo de desarrollo de capacidades para seleccionar la evaluación que se realizaría (encuentros, curso y convocatoria) como se relata en las páginas 8 y 9 del informe, desarrollándose entre junio de 2014 y junio de 2016. Junto a ello, y ya más en el ámbito de esta pregunta de metaevaluación, la evaluación no contempla un único taller de inicio, sino una secuencia de talleres dirigidas a los stakeholders, en este caso los ya implicados en el equipo de evaluación. El equipo de evaluación se reunió en 12 ocasiones/talleres entre julio de 2016 y diciembre de 2017. No se contemplaron al inicio otros talleres con stakeholders al margen de los comentados con el equipo evaluador.
		No hubo un taller de inicio con stakeholders para facilitar su participación	1	
		No hubo un taller de inicio con stakeholders pero fueron implicados bajo otra fórmula (aisladamente, presentaciones, etc.)	5	
		Hubo un taller de inicio con stakeholders	**10**	
2.5	¿Los informes de evaluación (o solo resúmenes o extractos) son accesibles para el público?	Ninguna / No aplica	0	El informe final de evaluación y el resumen ejecutivo está disponible tanto en la web de MIDEPLAN como en la de FOCEVAL: http://foceval.org/2017/02/presentacion-de-resultados-de-la-evaluacion-participativa-de-los-servicios-de-prevencion-y-atencion-del-cancer-costa-rica/ http://www.mideplan.go.cr/component/content/article?id=1695
		No son accesibles para el público (ni Informe ni resumen)	1	
		Son accesibles para el público (Resumen o extracto)	5	
		Son accesibles para el público (Informe y resumen)	**10**	
3.	**Credibilidad y exactitud**	**Calificación (texto)**	*Calificación (numérico)*	**Observaciones cualitativas**
3.1	**Metas, objetos y modelo de cambio**			
3.1.1	¿Está descrita la meta/el objetivo de la evaluación?	Ninguna / No aplica	0	En la página 20 se dedica un apartado a exponer con detalle los objetivos de la evaluación, tanto el objetivo general como su articulación a partir de tres objetivos específicos. El objetivo general que se planteó fue el de proponer mejoras a partir de la identificación y valoración de los aciertos y desaciertos de los servicios de prevención y atención del cáncer que presta la Caja en Valle de la Estrella. Son objetivos específicos, valorar: 1) si los servicios que se prestan en Valle de la Estrella se ajustan a las necesidades reales de sus pobladores; 2) cuál es la calidad de los servicios prestados y qué elementos se pueden mejorar; y 3) si la Caja está combinando de forma óptima los recursos materiales y humanos disponibles para prestar estos servicios.
		No se describe el objetivo de la evaluación (Ni TdR ni Informe)	1	
		Se describe el objetivo de la evaluación de manera imprecisa	5	
		Se describe de manera precisa el objetivo	**10**	

				4) mejorar la colaboración entre el personal de la Caja y las Juntas de Salud.
3.1.2	¿Se describe con precisión el objeto de la evaluación (p.ej. grupos interesados, dimensión regional, volumen financiero, duración del proyecto, niveles de intervención etc.)?	Ninguna / No aplica No se describe el objeto de evaluación Se describe de manera imprecisa el objeto evaluado (>50% de ítems contemplados) o no contempla aspectos relevantes **Se describe con precisión (al menos los ítems contemplados)**	0 1 5 **10**	El informe dedica un apartado a descripción del objeto evlauado ("El programa evaluado", Pág. 29) donde se explica el detalle del funcionamiento de los servicios que se evalúan. Se detalla el marco de referencia de planificación (Plan Nacional para la Prevención y Control del Cáncer 2012 – 2017 (Ministerio de Salud, 2012)), su propósito, estrategia, así como sus ámbitos de actuación sobre los que se focaliza esta evaluación (delimitación del objeto). Se describe describe el proceso que conforman los servicios de prevención y atención del cáncer, a partir de la reconstrucción realizada. Al hilo de esta descripción se detallan algunos grupos o agentes implicados en algunas de sus etapas. No obstante, respecto a los ítems que contempla esta pregunta de metaevaluación cabe apuntar que en el informe no se hacen referencias al volumen financiero (global o en la zona de actuación), ni a la duración del programa (o el horizonte temporal analizado). La referencia temporal no es algo relevante para esta evaluación y el volumen financiera es habitual que no sea refereido en las evaluaciones metaevaluadas, adoptando definiciones de eficiencia vinculadas al uso de otros recursos distintos al económico-financiero.
3.1.3	¿Se realizó un análisis de los grupos de interés (stakeholder)?	Ninguna / No aplica No se realizó un análisis de stakeholders **Se realizó someramente o no se aprecia su uso (técnicas, recomendaciones, etc.)** Se realizó y fue utilizado como eje de la evaluación (y epígrafe explícito)	0 1 **5** 10	En el informe de esta evaluación no se realiza ningún análisis explícito de stakeholders. A lo largo del informe si se tienen en cuenta una serie de perfiles destinatarios, sobre todo cuando se plantean las actividades de cara a su socialización y cuando se sistematizan las recomendaciones de mejora, aunque también cuando se exponen las fuentes consultadas. También la graficación del proceso del servicio que se presta de prevención y atención del Cáncer así como en algunas referencias a lo largo del informe emergen algunos de los principales stakeholders. En tanto que la división "prestador del servicio-usuario-comunidad" es un eje que se mantiene a lo largo del informe se entiende que se ha realizado, pero no se ha explicitado en el informe y resulta aventurado juzgar su sistematicidad. No aparece una reflexión que ayude a comprender si este análisis se tuvo en cuenta implícitamente al menos al inicio para determinar la configuración del equipo de evaluación. Si en la delimitación del objeto, se detalla por ejemplo que hay algunos ámbitos del plan nacional que se priorizan (entendiéndose que otros no se atienden), o en el análisis de eficacia se descarta priorizar algunas preguntas, este tipo de análisis no se aprecia en el caso de los stakeholder. Son estos motivos los que llevan a otorgar una calificación intermedia en este ítem, no tanto por el carácter somero del análisis (como apunta el ítem) sino por falta de explicitación y sistematicidad del mismo en cuanto al informe se refiere.
3.1.4	¿Se elaboró o se usa una concluyente teoría de la intervención, cadena de resultados o lógica de impacto?	Ninguna / No aplica No se elaboró Se elaboró pero no se usa en la evaluación o no es concluyente **Se elaboró, se usa en la evaluación y es concluyente**	0 1 5 **10**	En la página 25 se indica que "se elaboró lo que en evaluación se llama teoría de la intervención, que no es sino una descripción del servicio o programa que se va a evaluar… Finalmente estas actividades permitieron ir delimitando el objeto de la evaluación, es decir los componentes del servicio de prevención y atención del cáncer que serían evaluados". Y, entre las páginas 29 a 34 se expone la delimitación de este objeto y gráfico de proceso que dibuyan sus actuaciones y niveles de intervención. Aunque cabe esperar que la teoría de la intervención incorpore la respuesta que se espera que el programa genere en los participantes, puede afirmarse que cuando se evalúan servicios y no programas (como es el caso), lo habitual es que se recurra a estos gráficos de proceso y a su explicación detallada como eje vertebrador. En este sentido, el carácter concluyente puede entenderse en la medida en que se aprecie el ciclo completo del servicio que permita juzgarlo de manera completa. De esta manera, queda explícita su elaboración en esta evaluación y se puede afirmar que ha sido utilizada a lo largo del informe, especialmente en la exposición de hallazgos y en los instrumentos aplicados (esta graficación permite ubicar con claridad los diferentes discursos de los participantes en cada etapa del proceso de prestación de este servicio globalmente considerado). Se aprecia además una concentración del foco de la evaluación en la prevención y en la atención primaria (quizás los mas cercanos al contexto local), no sólo en los discursos de los usuarios (más en antención primaria) sino también en los stakeholders considerados. No obstante, hubiese sido deseable que la exposición de hallazgos y conclusiones hubiese facilitado esta lectura procesual en mayor medida, pues requiere la lectura completa del informe y de sus anexos.
	¿Se formularon hipótesis de impacto concluyentes?	**Ninguna / No aplica** No se formularon	**0** 1	

		Se formularon pero no son concluyentes	5	No se recoge dentro del Informe final ninguna hipótesis formulada de impacto relativa al programa evaluado, ni tan siquiera de eficacia, pues ubica el foco de análisis de esta última en la calidad del servicio por las propias limitaciones de su diseño, como se explicita (Pág. 39). Esta situación es coherente en tanto que se trata de una evaluación de la calidad de un servicio público y no del impacto del mismo, que no forma parte de las necesidades de información que muestran sus preguntas y no es, por tanto, un criterio evaluativo. Por ello, se entiende que este ítem no aplica a esta evaluación.
		Se formularon y son concluyentes	10	
	¿Se diferencia entre los diferentes niveles de impacto? Niveles de impacto:	**Ninguna / No aplica**	**0**	En el Informe final no se diferencia en ningún momento entre distintos niveles de impacto. Tal y como se indica en la observación anterior, esto puede ser coherente en tanto que se trata de una evaluación de la de un servicio público entre cuyos criterios está la calidad pero no el impacto. Por ello, se entiende que este ítem no aplica a esta evaluación.
		No se diferencia entre niveles de impacto	1	
		Se diferencia algo (impactos intermedios/finales o dimensiones en las que opera el impacto)	5	
		Se diferencia con claridad los diferentes niveles de impacto	10	
3.1.5	¿Se describe de forma exhaustiva el contexto de la evaluación y se relaciona éste con la teoría de intervención del proyecto?	Ninguna / No aplica	0	Este ítem se valora en parte en los informes a través del capítulo que suelen dedicar a este fin, exponiendo las particularidades del contexto evaluado, normalmente con referentes de ámbitos superiores (regional, nacional, etc.) que permitan ubicar al lector y que permitan apreciar igualmente las decisines de la evalaución en su contexto. El informe de evalaución describe someramente algunos datos y rasgos caracterizadores en su presentación pero no puede decirse que realice una presentación diferenciada y exhaustiva de su contexto evaluativo. A lo largo del informe, con algunos hallazgos (con fuente estadística) y con los propios testimonios de los participantes (especialmente los que se relatan en su anexo), el lector va aproximándose al contexto. El hecho de que se trate de una evaluación participativa en la que sus protagonistas son conocedores y agentes clave del entorno, y sus destinatarios principales también (ellos mismos de alguna manera), puede justificar que el informe preste poca atención a esta necesaria contextualización, pues no debiera mermar su utilidad para quienes evalúan y que son en parte receptores de la misma, si bien limitan la bondad del propio informe para otras lecturas o lectores. Más que relacionar el contexto con la teoría de la intervención se puede decir que son los hallazgos los que tratan de vincularse con dicha teoría.
		No se describe el contexto de la evaluación	1	
		Se describe el contexto no exhaustivamente o no se relaciona con la teoría de la intervención	**5**	
		Se describe exhaustivamente el contexto y se relaciona con la teoría de la intervención	10	
3.2	**Diseño y métodos**			
3.2.1	¿Se corresponde el diseño de la evaluación con los intereses de la misma?	Ninguna / No aplica (o los TdR no explicitan los intereses de la evaluación)	0	No se dispone de TdR para valorar los intereses de la evaluación, por lo que dichos intereses son contemplados a paritr de lo que expone el propio informe. El diseño responde a una evaluación participativa, que se ubica entre los "Enfoques Colaborativos de Evaluación" como el propio informe indica (Pág.22). Sus rasgos definitorios, funciones y bondades son tratadas en profundidad y con un marcado carácter pedagógico a lo largo del informe. Antes de juzgar su adecuación a los intereses de la evalaución como contempla este ítem es preciso apuntar que si bien es deseable que "la pregunta preceda al método", en este caso la propia convocatoria determinaba ya la elección de una evaluación participativa. Esto que no es demérito, hace que ya las propuestas presentadas muestren al menos unos objetivos e intereses que debieran satisfacerse para los promotores de la misma a través de esta opción de diseño. Dicho esto, y entrando ya en el propio diseño de la evaluación participativa, a tenor del proceso desplegado para la formulación de preguntas se puede afirmar que este diseño responde a los intereses que la motivan, tanto sus objetivos, general y específicos, como sus preguntas y criterios evaluativos. La vertebración de las preguntas a través de los criterios de valor es un elemento que facilita o ayuda a valorar la bondad de este diseño pues no todas las evalauciones participativas plantean criterios de esta naturaleza (evaluaciones comprensivas o evaluaciones sin indicadores). Esto ha facilitado que la evaluación no haya eludido emitir juicios de valor sobre los criterios que la orientaban (pertinencia, eficacia y eficiencia) así como dar respuesta suficiente a las preguntas. En este sentido, el diseño también se desarrolla dentro de unos márgenes en tanto que busca alternativas para ofrecer las mejores respuestas reconociendo explícitamente sus limitaciones para atender otras necesidades informativas. Este hecho, implícito en la evaluación participativa, hace que no que no queden preguntas sin respuesta sino que el equipo evalaudor vaya moldeando aquellas siendo conscientes de aquello sobre lo que puede obtener respuesta.
		El diseño no se corresponde con los intereses explícitos de la evaluación (según los TdR o el propio Informe)	1	
		El diseño se corresponde con los intereses implícitos de la evaluación o se corresponde moderadamente con los objetivos explícitos (según los TdR o el propio Informe)	5	
		Se corresponde el diseño de la evaluación con sus intereses explicitados en los TdR (según los TdR o el propio Informe)	**10**	
3.2.2		Ninguna / No aplica	0	

	¿Se explican y justifican suficientemente el diseño y los métodos aplicados en la evaluación?	No se explicitan ni justifican el diseño y los métodos aplicados	1	La elección de la evaluación participativa se encuentra suficientemente explicada y justificada. En el caso de los métodos y técnicas empleadas para su desarrollo (entrevistas, grupos focales, historias de vida) su justificación se encuentran en los mismos anexos en los que se presentan. Se expone también la triangulación como estrategia adoptada y se relata la sistematicidad con que se ha aplicado (agrupando hallazgos de cada perfil y técnica según pregutnas de evaluación en una matriz en Excel). En el informe los hallazgos se presentan ya integrados en función del colectivo del que provienen y no desde la técnica de la que provienen.
		Se explicitan pero no se justifican suficientemente el diseño y los métodos aplicados	5	
		Se explicitan y justifican suficientemente el diseño y los métodos aplicados	**10**	
3.2.3	¿Se consideró lo contrafactual? ¿Cómo? (p.ej. ¿se construyeron grupos de comparación con posterioridad?)	**Ninguna / No aplica**	**0**	No se consideró lo contrafactual ni la construcción de grupos de comparación y puede afirmarse que no aplican a una evaluación participativa como la que aquí se plantea.
		No se consideró lo contrafactual	1	
		Se consideró tangencialmente lo contrafactual (aproximación) o no se especifica suficientemente cómo se consideró	5	
		Se consideró lo contrafactual y se especificó cómo	10	
3.2.4	¿Se aplicaron diferentes métodos de levantamiento de datos, cubriendo aspectos cuantitativos y cualitativos?	Ninguna / No aplica	0	La evaluación participativa contempló métodos para el levantamiento de datos que cubren aspectos cualitativos. Los aspectos cuantitativos presentan un carácter testimonial en esta evaluación y, al margen de los que puedan contemplar alguna de las fuentes consultadas, en el informe prácticamente se concentran en unos datos sobre evolución de la proporción de enfermos de cáncer fallecidos cada año (gráfico en Pág.43).
		Los métodos aplicados no cubren aspectos cuanti ni cuali	1	
		Los métodos aplicados cubren aspectos cuanti o cuali	**5**	
		Los métodos aplicados cubren aspectos cuanti y cuali	10	
uipo	¿Qué métodos de recogida de datos fueron aplicados y a cuántas personas se consultó a través de cada uno (p.ej. número de entrevistas guiadas, o número de personas entrevistadas online)? a) Entrevistas guiadas (número?) ; b) grupos focales (número?) c) encuestas (semi-) estructuradas (número?) ¿de qué forma (face-to-face, online, teléfono etc.) ; d) otros?...	Ninguna / No aplica	0	
		Se aplicó sólo un método	1	
		Se aplicaron dos métodos	5	
		Se aplicaron una pluralidad de métodos (más de dos)	**10**	

Técnica / Método	Tipo	Grupo de actores	Participantes
Entrevistas	Semiestructuradas	Profesionales sector salud, Personas diagnosticadas con cáncer o familiares	4 5
Grupos	Focales	Profesionales del sector salud: Enlace de Oncología Limón Personal médico del Valle de la Estrella	9 7
Historias de vida	Individual	Personas usuarias del servicio de salud	8
	Colectiva	Personas usuarias del servicio de salud	1
Juego de simulación	"Mitos y creencias sobre el cáncer"	Personas de la comunidades de San Rafael y la Guaria (Valle de la Estrella)	13
		Total	47

3.3	**Procedimiento de selección**			
3.3.1	¿Se explicita el procedimiento y los criterios de selección en el informe?	Ninguna / No aplica	0	El procedimiento y los criterios de selección aparecen especificados para las técnicas aplicadas en los anexos del informe, especialmente detallados para los grupos focales (Pág.74), con menor detalle para las entrevistas (perfiles concretos) o las historias de vida (selección de comunidades) procedimiento para cumplimentar los instrumentos de recogida de formación cuantitativa 1 y 2, como la determinación de la población meta a partir de la que extraer la muestra objeto de estudio. (Páginas 27-31) (ver Tablas Nº3 y Nº4) Dentro del epígrafe 3.7.2 Enfoque cualitativo, se especifican los criterios de selección de la población participante en los Grupos focales, así como los responsables de efectuar dicha selección. (Página 32-33)
		No se explicitan los procedimientos ni criterios de selección	1	
		Se explicitan en algunos casos (técnicas)	5	
		Se explicitan los procedimientos y criterios de selección	**10**	
3.3.2	¿Qué procedimiento de selección fue aplicado y que tan extensas son las muestras? Se entiende que aplica esta pregunta para las Encuestas	**Ninguna / No aplica**	**0**	En este ítem de la metaevaluación la mirada se ha concentrado en las encuestas contempladas en las evaluaciones, aunque no de manera exclusiva. Por ello, su valoración suele estar muy influenciada por estas encuestas, considerándose que no aplica en aquellos casos en los que no hay encuesta. No obstante, sobre este extremo, en esta evaluación se apunta lo siguiente: "La cantidad de personas entrevistadas o participantes de los talleres y grupos focales se determinó en función del alcance de la evaluación y el tiempo disponible. No responden a ningún criterio de representatividad estadística aunque,
			1	
		La muestra no es adecuada o el error no es aceptable	5	
		La muestra sí es adecuada y el error aceptable	10	

				en conjunto, intentan reflejar una valoración global del servicio desde distintas perspectivas e incorporar elementos del contexto particular de las personas que participaron. Las conclusiones y resultados obtenidos con esta evaluación son, por lo tanto, de tipo cualitativo e interpretativo" (Página 26).
3.4	**Análisis de datos**			
3.4.1	¿Se triangulan datos cuantitativos y cualitativos (métodos mixtos)?	Ninguna / No aplica	0	Esta evaluación utiliza la triangulación como estrategia de análisis, como apunta el propio informe. La triangulación se realiza entre los datos cualitativos recabados a través de diversas técnicas y se hace de manera generalizada, pero no puede decirse que esta triangulación se efectúe entre datos cuanti y cuali (los datos cuantitativos son la excepción en esta evaluación..
		No se triangulan datos cuanti-cuali	1	
		Se triangulan datos cuanti-cuali pero no de manera generalizada	**5**	
		Se triangulan datos cuanti-cuali de manera generalizada	10	
3.4.2	¿Se cuantifican la respuesta siempre y cuando esto sea posible (p.ej. 60% de los entrevistados, o tres de 5 personas entrevistados…?	Ninguna / No aplica	0	La excepción en el informe es que se cuantifiquen las respuestas obtenidas a través de las técnicas aplicadas. La naturaleza cualitativa e interpretativa de la evaluación y también su alcance en cuanto personas implicadas en cada una de estas técnicas, hace que la cuantificación de respuestas no sea un recurso utilizado. Las excepciones quizás se concentran en la página 41, donde se usa este recurso en relación con las entrevistas y con un hallazgo sobre el juego reaizado (en este juego sí se contempla la cuantificación en lo posible –Anexo-).
		No se cuantifican las respuestas cuando es posible	1	
		Se cuantifican las respuestas a veces cuando es posible	**5**	
		Se cuantifican las respuestas de manera generalizada cuando es posible	10	
3.4.3	¿Se usaron datos del monitoreo del proyecto/programa?	Ninguna / No aplica	0	Aunque entre las fuentes consultadas aparecen memorias institucionales del Ministerio de Salud no se aprecia que la evaluación haya usado datos del monitoreo del programa (podría considerarse nuevamente en este caso que la excepción serían los datos sobre fallecimientos –pág. 43-).
		No se usaron datos del monitoreo	**1**	
		Se usaron datos del monitoreo de manera insuficiente	5	
		Se usaron datos del monitoreo de manera suficiente	10	
3.4.4	¿Se recurrió a fuentes de datos externos (p.ej. estadísticas) para el análisis?	Ninguna / No aplica	0	Las fuentes externas contempladas se utilizan para ofrecer un par de datos en la página de presentación a modo de contextualización (Pág.7) y para mostrar un hallazgo en la evaluación de la eficacia (Pág.43). Cabe apuntar que hubiese sido posible ofrecer una contextualización más detallada así como utilizar fuentes externas para reforzar la triangulación (pertinencia y eficiencia). Es este un recurso que usado moderadamente no debería entrar en conflicto con los requerimientos de la evaluación participativa.
		No se recurrió a fuentes externas	1	
		Se recurrió a fuentes externas de manera insuficiente	**5**	
		Se recurrió a fuentes externas de manera suficiente	10	
3.5	**Criterios de evaluación**			
3.5.1	¿El informe de evaluación expone de manera concluyente y comprensible, si y hasta qué punto el proyecto/programa			Esta evaluación establece como criterios la eficacia (calidad), eficiencia y pertinencia.
	a) es relevante/pertinente para el país, la región y el grupo meta	Ninguna / No aplica	0	Con pocos datos de contexto el informe muestra la relevancia de este servicio para el país. En cuanto a la pertinencia el informe ofrece hallazgos relevantes en las diferentes dimensiones consideradas (ajuste a expectativas y necesidades; accesibilidad y coherencia con el PNC). Aunque se haya contemplado la exposición de hallazgos como respuestas a las preguntas planteadas/objetivos específicos (como se apunta en la Pág.58), y esto haya diluido en parte el juicio evaluativo, se expone con claridad la conclusión sobre la necesidad de mejorar el ajuste a la realidad y necesidades de la población, y lo hace de manera comprensible.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible, o es comprensible pero no concluyente	5	
		Lo expone de forma concluyente y comprensible	**10**	
	b) logró las metas definidas	Ninguna / No aplica	0	La evalaución concibe la eficacia más como componente de calidad del servicio que como logro de metas en sí. Contemplar la exposición de hallazgos como respuestas a las preguntas planteadas/objetivos específicos (como se apunta en la Pág.58), ha derivado en que el juicio evaluativo haya quedado diluido, y se aprecia también un marcado carácter de síntesis o de balance de hallazgos más que la emisión de un juicio evaluativo concluyente en lo que se presenta como conclusión sobre eficacia/calidad, aunque si es comprensible. En este caso se encuentra como factor añadido que no se identificaron dimensiones clave de calidad que ayudasen a vertebrar las respuestas como sí ocurriera en el caso de la pertinencia (adecuación, accesibilidad, coherencia),
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	
		Lo expone de forma concluyente y comprensible	10	

				Además, se contemplan aspectos diversos bajo este criterio, como el grado de conocimiento que si puede considerarse un logro de la prevención o información recibida. En este caso, el juego realizado ha sido un recurso que ha ocupado poco espacio en los hallazgos, y que quizás podría haber aportado mayor y más detalllada información sobre alguno de los indicadores planteados que no encontraron respuesta (como el relativo a "Nivel de conocimiento sobre cada tipo de cáncer por parte de pobladores y/o usuarios"). Su aplicación en dos grupos reducidos (uno de ellos especialmente, 3) hace que finalmente se presente más como un suproducto de la evaluación, muy útil para los procesos de comunicación, que como una técnica que permita ponderar el siquiera cualitativamente el grado de conocimiento, lo que hubiese requerido un mayor alcance de esta técnica atendiendo a los varios indicadores que en la matriz atendían a esta cuestión de la divulgación y el conocimiento. O incluso para ofrecer pistas sobre en qué medida dichos mitos o creencias se encuentran asociados en atención que presta o no presta la población a la propia labor de prevención. En lo que se presenta como conclusiones de este criterio (Pág.58) se aprecia más un carácter de síntesis de hallazgos que de emisión de un juicio evaluativo. En este sentido llama también la atención que las conclusiones no se formulen y expongan de manera previa a las recomendaciones.
	c) es eficiente	Ninguna / No aplica	0	La lectura del cuadro de hallazgos (Pág.44) no encuentra una exposición concluyente sobre la eficiencia (Pág.58), donde se aprecia más un carácter de síntesis de algunos hallazgos que de emisión de un juicio evaluativo. En este sentido llama también la atención que las conclusiones no se formulen y expongan de manera previa a las recomendaciones.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	**5**	
		Lo expone de forma concluyente y comprensible	10	
	d) causó impactos (outcome e impacto) deseados y no deseados.	**Ninguna / No aplica**	**0**	El impacto no es un criterio de esta evaluación ni tiene preguntas que responder en este sentido.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	e) causó impactos duraderos	**Ninguna / No aplica**	**0**	El impacto no es un criterio de esta evaluación ni tiene preguntas que responder en este sentido.
		No lo expone	1	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
	f) cuenta con un sistema de monitoreo y evaluación y si éste fue usado para la gestión (calidad de la gestión del proyecto)	Ninguna / No aplica	0	
		No lo expone	**1**	
		Lo expone pero no de forma concluyente y comprensible	5	
		Lo expone de forma concluyente y comprensible	10	
3.5.2	g) ¿Los cambios observados se reconducen al proyecto/programa de manera concluyente?	Ninguna / No aplica	0	Entre los principales cambios observados que se apuntan cabría destacar el grado de conocimiento de la población y a la incidencia del cáncer y las defunciones. Si bien en el caso del conocimiento de la población se puede decir que se reconduce al proyecto, su limitada aplicación no permite un carácter concluyente. Se introduce un hallazgo sobre incidencia del cáncer/mortalidad que quizás hubiese requerido mayor tratamiento o información complementaria para que el lector pueda formarse una opinión al respecto. En cualquier caso no se reconduce al programa/servicio de manera concluyente.
		No se identifican cambios observados o no se reconducen al proyecto de manera concluyente	1	
		Se identifican cambios observados pero no se reconducen al proyecto/programa de manera concluyente.	**5**	
		Se identifican cambios observados y se reconducen al proyecto/programa de manera concluyente.	10	
4.	**Calidad del informe y utilidad**	Calificación (texto)	*Calificación (numérico)*	Observaciones cualitativas
4.1		No aplica: No se dispone de los TdR o no se indican expresamente las preguntas en ellos formuladas	0	Aunque no se dispone de TdR, el anexo muestra la matriz de evaluación y el detalle de preguntas, que también se exponen con detalle en el informe.

	¿El informe de evaluación responde de forma adecuada a todas las preguntas listadas en los TdR?	No responde adecuadamente a ninguna pregunta	1	Se puede afirmar que el informe ofrece una respuesta adecuada a las preguntas de evaluación formuladas, especialmente a las que se indican en el propio texto del informe.
		No responde adecuadamente a alguna/s pregunta/s	5	
		Responde adecuadamente a todas las preguntas	**10**	
4.2	¿Se describen de forma balanceada los puntos fuertes y débiles del programa/proyecto?	Ninguna / No aplica	0	Los cuadros de hallazgos que se presentan son una buena ilustración de una lectura balanceada de puntos fuertes y débiles, aunque en algunos casos puedan destacar puntos fuertes o débiles según la percepción de los usuarios o profesionales (como ocurre en la eficiencia sobre la adecuación y suficiencia de recursos)
		No se describen puntos fuertes y débiles	1	
		Se describen principalmente puntos fuertes o débiles	5	
		Se describen balanceadamente puntos fuertes y débiles	**10**	
4.3	¿Se distingue entre declaraciones descriptivas y declaraciones valorativas?	Ninguna / No aplica	0	El informe es claro en este sentido, mostrando con claridad cuando se realiza alguna interpretación o valoración sobre los hallazgos o simplemente cuando se están exponiendo declaraciones descriptivas.
		No se distinguen con claridad declaraciones descriptivas y valorativas	1	
		Se distinguen declaraciones descriptivas y valorativas en algunos apartados del informe	5	
		Se distinguen declaraciones descriptivas y valorativas en el informe en su conjunto	**10**	
4.4	¿Se especifica como las estimaciones y valoraciones provienen de datos primarios y secundarios confiables?	Ninguna / No aplica	0	Las valoraciones se asocian con claridad a los datos de los que provienen y la confiabilidad de los datos/hechos/percepciones recabados ha quedado expuesta en el informe para los fines de esta evaluación participativa de naturaleza cualitativa e interpretativa, en la que las muestras no persiguen representatividad sino facilitar la comprensión de los fenómenos estudiados desde diversos puntos de vista. En el caso de que exista alguna limitación en la interpretación que se puede efectuar de los hallazgos esto también se expone de manera clara.
		No se especifica	1	
		Se especifica en ocasiones	5	
		Se especifica en general	**10**	
4.5	¿Las conclusiones están basadas en los resultados?	Ninguna / No aplica	0	Las conclusiones se vinculan a los hallazgos obtenidos y están ancladas en ellos. Esto es más evidente si se tiene en cuenta que se toman los hallazgos como respuesta a las preguntas y las conclusiones constituyen una síntesis de hallazgos/resultados.
		Las conclusiones no están basadas en los resultados	1	
		Alguna o algunas conclusiones no están basadas en resultados	5	
		Todas las conclusiones están basadas en los resultados	**10**	
4.6	¿Se llega de forma lógica y comprensible de las recomendaciones a las conclusiones?	Ninguna / No aplica	0	Las recomendaciones permiten apreciar las conclusiones/hallazgos de los que se derivan, pues son esencialmente los hallazgos el eje vertebrador en esta evaluación.
		No se llega ni de forma lógica ni comprensible	1	
		Se llega de forma lógica o comprensible pero no las dos	5	
		Se llega de forma lógica y comprensible	**10**	
4.7	¿Las recomendaciones son relevantes, específicas para los destinatarios y concretas?	Ninguna / No aplica	0	Se considera que las recomendaciones son relevantes para la mejora del programa y su especificidad radica en que los evaluadores han negociado/consensuado unas recomendaciones muy focalizadas en aquello que se encuentra en su entorno más próximo y donde se espera que su capacidad de influencia como gestores (profesionales médicos, enlaces y coordinadores de oncología) y usuarios (Junta deSalud) pueda inducir mejoras efectivas.
		No son relevantes ni específicas	1	
		Son relevantes o específicas pero no las dos	5	
		Son relevantes y específicas	**10**	
4.8	¿El informe final contiene un resumen estructurado, bien legible y comprensible?	Ninguna / No aplica	0	El resumen se acompaña de un resumen ejecutivo bien estructurado, en el que se presentan de forma esquemática los aspectos de los que trata la evaluación. El resumen no expone las principales conclusiones y hallazgos, mencionándose sólo las recomendaciones.
		El Informe no contiene un resumen	1	
		El informe contiene un resumen estructurado	**5**	
		El informe contiene un resumen estructurado, bien legible/comprensible	10	
5.	**Exhaustividad y adecuación de la documentación**			
5.1		**Ninguna / No aplica**	**0**	No se formularon los TdR de esta evaluación..
		Los TdR no son accesibles en anexos	1	

	¿Los TdR son accesibles en el anexo en la lengua común de trabajo o en una traducción?	Los TdR son accesibles en anexos pero no en la lengua común de trabajo o en una traducción	5	
		Los TdR son accesibles en la lengua común de trabajo o en una traducción	10	
5.2	¿Se incluyen en los anexos el cronograma de la fase de campo, incluyendo listados de lugares y organizaciones visitadas y de personas entrevistadas?	Ninguna / No aplica	0	Se incluye un cronograma de la evaluación que complmente a la información que ya aparece en el cuerpo del informe (Página 9) y en el anexo 3 dedicado a las fuentes de información incorpora suele incorporar las fechas y lugares de aplicación de cada una de ellas, salvo en las entrevistas. No se registra el nombre de las personas participantes en el trabajo de campo ni tampoco es habitual que se recojan las organizaciones participantes y personas entrevistadas, que si bien en las dinámicas grupales e historias de visa puede ser comprensible, tampoco los muestra en el caso de las entrevistas. .
		No se incluyen en anexos	1	
		Se incluyen algunos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas) en anexos	**5**	
		Se incluyen en anexos (cronograma fase de campo y listados de lugares, organizaciones y personas entrevistadas)	10	
5.3	¿Se incluye en los anexos la documentación de las guías para grupos focales, entrevistas guiadas y encuestas, así como una lista de la literatura y documentación utilizada?	Ninguna / No aplica	0	En los anexos se incluyen los documentos para la recogida de información (Anexo 3) y el informe también incluye las referencias consultadas.
		No se incluyen en anexos (guías y fuentes)	1	
		Se incluyen algunos en anexos (guías o fuentes)	5	
		Se incluyen en anexos (guías y fuentes)	**10**	
5.4	¿El informe está estructurado de forma adecuada y su lectura se ve facilitada con gráficos, tablas y resúmenes de los capítulos? **Elementos que faciliten su lectura**: gráficos, tablas y resúmenes de capítulos.	Ninguna / No aplica	0	El informe final de evaluación presenta una estructura correcta para facilitar su lectura. La propia redacción de los epígrafes, asociando una pregunta o un enunciado claro y directo dirigido al lector es un elemento que ayuda a su lectura. La redacción es clara y estilo es netamente pedagógico, procurando que todo aquel aspecto técnico de evaluación que se pueda exponer quede suficiente y sencillamente explicado. Es habitual la utilización de gráficos que van a ayudando la lectura y comprensión y también tablas y gráficos, aunque estos últimos en mucha menor medida; complementando eso con fotos del equipo evalaudor y de sus sesiones de trabajo junto a otras. También su utilizar párrafos para ir recapitulando aquello que se explicando y exponiendo, aunque no se puede decir que constituyan en todos los casos un resumen como tal de los capítulos o apartados. Además se acompaña de un video que facilita la labor de comunicación del proceso.
		El informe no presenta una estructura adecuada ni elementos que faciliten su lectura	1	
		El informe presenta una estructura adecuada o elementos que facilitan su lectura; o, cubriendo ambas facetas, su registro en alguna de ellas es moderado (estructura/ elementos facilitadores).	5	
		El informe presenta una estructura adecuada y elementos que facilitan su lectura (gráficos, tablas y resúmenes de capítulos)	**10**	