

eman ta zabal zazu



Universidad del País Vasco Euskal Herriko Unibertsitatea

**Departamento de Economía Aplicada I**

**Tesis doctoral**

**La capacidad evaluativa en el Sistema de las Naciones Unidas**

**Programa de Doctorado en Globalización, desarrollo y cooperación internacional**

**Presentada por: Itziar Arispe Ruiz de Gauna**

**Dirigida por: María Ángeles Díez López**

**Noviembre 2015**









## *Agradecimientos*

A lo largo de estos años han sido muchas las personas que me han brindado su apoyo y colaboración para la realización de esta tesis y a quienes deseo expresar mi más sincero agradecimiento.

En primer lugar, debo un especial reconocimiento a mi Directora de tesis, Maria Ángeles Díez, por su guía, apoyo y siempre oportunos consejos sobre la orientación de la tesis así como por su pensamiento crítico. Agradezco igualmente su trato cercano y plena disponibilidad a pesar de que hayamos tenido que trabajar desde la distancia.

También quiero extender mi agradecimiento al Instituto de Estudios sobre Desarrollo y Cooperación Internacional, Hegoa, así como al Departamento de Economía Aplicada I de la Facultad de Ciencias Económicas y Empresariales de Bilbao (Sarriko) por su 'acogida' en su Programa de Doctorado y por el apoyo recibido durante todo el periodo de investigación.

Me gustaría dar las gracias también a Jonas Buche y a Markus Siewert de la Universidad Goethe de Frankfurt, por su generosidad al prestarme su apoyo y buenos consejos en relación al método principal de investigación que he empleado en este trabajo (Análisis Comparativo Cualitativo). Su buen humor y cordialidad hicieron de mi visita a la Universidad de Tilburg, no solo una experiencia de gran utilidad sino también muy grata.

A lo largo de todos estos años en los que compaginar mi trabajo con la labor investigadora no siempre ha sido fácil, no podía sino trasladar un sincero y especial agradecimiento a mis compañeros y compañeras con los que he tenido la oportunidad de trabajar tanto en el Organismo Internacional de la Energía Atómica (OIEA) como en el Instituto de Investigación y Capacitación Interregional sobre Crimen y Justicia (UNICRI) y en especial a evaluadores y evaluadoras como Frits Eriksen, Fumika Ouchi, Dalibor Kysela, Oscar Hernández y Roger Miranda, con los que he compartido incontables horas de oficina, misiones de evaluación y vivos debates. También quiero agradecer a todos aquellos con los que he interactuado a lo largo de estos años en el marco de las evaluaciones de programas y proyectos que los que he tenido la oportunidad de participar, por sus valoraciones y críticas, que aunque severas en algunas ocasiones, todas ellas han contribuido a forjar mi pasión por esta profesión.

Finalmente, todo esto nunca hubiera sido posible sin el apoyo incondicional de mi familia, mis padres Pilar y José Luis, a quienes les debo eternamente su aliento e inspiración, a Feli y a mi hermano Javier, ejemplos de superación, a Carmen y a Josu, grandes entusiastas del mundo educativo y, por supuesto, a mi compañero en éste y en tantos otros viajes, Asier, así como a nuestros 'pequeños investigadores', Eneko y Ander. Gracias por existir.



# ÍNDICE

<b>Glosario .....</b>	<b>1</b>
<b>Introducción .....</b>	<b>5</b>
<b>1. El valor de la evaluación desde su utilidad .....</b>	<b>13</b>
1.1. La evaluación como concepto general .....	13
1.2. La utilización en evaluación.....	24
1.3. Los tipos de uso de las evaluaciones .....	30
1.4. El mal uso en la práctica evaluativa.....	35
1.5. La investigación empírica existente sobre la utilización .....	37
1.6. Debates conceptuales recientes: de la utilización a la influencia .....	42
1.7. La evaluación y su convivencia con otras fuentes de información .....	48
<b>2. La capacidad evaluativa de las organizaciones .....</b>	<b>61</b>
2.1. El concepto de capacidad evaluativa.....	61
2.2. La capacidad evaluativa en el ámbito de la ayuda oficial al desarrollo .....	64
2.3. Evaluación y capacidad evaluativa .....	69
2.4. La influencia teórica del enfoque participativo en la capacidad evaluativa.....	71
2.4.1 La evaluación participativa práctica y la evaluación participativa transformativa.....	73
2.4.2 La evaluación centrada en la utilización .....	76
2.4.3 La Evaluación Empoderativa .....	81
2.4.4 La Evaluación Colaborativa .....	88
2.5. Los modelos de capacidad evaluativa .....	92
2.5.1 El Modelo Multidisciplinario .....	92
2.5.2 El Enfoque de Inmersión Colaborativa .....	100
2.5.3 El Marco Interactivo Cultural.....	102
2.5.4 El Modelo Integrador de Capacidad Evaluativa .....	106
<b>3. La evaluación en el Sistema de las Naciones Unidas.....</b>	<b>111</b>
3.1. El Sistema de las NU: la función de evaluación y el objeto evaluado .....	112
3.2. Formas de evaluación en el Sistema de las NU .....	115
3.2.1 Dependencia Común de Inspección (Joint Inspection Unit, JIU).....	115
3.2.2 La Oficina de Servicios de Supervisión Interna (OIOS) .....	117
3.2.3 La evaluación en los organismos de Secretaría General.....	119
3.2.4 La evaluación en los Programas y Fondos.....	121
3.2.5 La evaluación en organismos especializados.....	122
3.2.6 La evaluación en organizaciones relacionadas.....	123
3.3. El marco normativo sobre evaluación en NU.....	124
3.3.1 El rol de la evaluación en el ciclo programático .....	124

3.3.2	Las Normas y Estándares UNEG .....	126
3.4.	La capacidad evaluativa en las NU.....	131
3.4.1	Principios rectores de la capacidad evaluativa .....	131
3.4.2	Instrumentos de análisis de la función de evaluación .....	133
<b>4.</b>	<b>Marco de análisis de la capacidad evaluativa .....</b>	<b>139</b>
4.1.	Condiciones de la capacidad evaluativa .....	139
4.1.1	Institucionalización de la función de evaluación.....	140
4.1.2	Políticas y prácticas evaluativas .....	142
4.1.3	Cultura de aprendizaje .....	144
4.2.	La hipótesis y su marco de análisis en el Sistema de las NU .....	146
4.3.	Metodología propuesta.....	148
4.4.	Las entidades objeto de estudio .....	162
<b>5.</b>	<b>Análisis empírico: Resultados y discusión .....</b>	<b>163</b>
5.1.	La calidad de la evaluación .....	163
5.1.1	Análisis de las condiciones necesarias.....	163
5.1.2	Análisis de las condiciones suficientes .....	170
5.2.	El uso sostenible de la evaluación.....	183
5.2.1	Análisis de las condiciones necesarias.....	183
5.2.2	Análisis de las condiciones suficientes .....	187
<b>6.</b>	<b>Conclusiones finales.....</b>	<b>197</b>
	<b>Bibliografía .....</b>	<b>209</b>
	<b>Anexos.....</b>	<b>231</b>
	<b>Anexo 1: Las entidades objeto de estudio .....</b>	<b>233</b>
1.1	UNO Habitat.....	233
1.2	PNUD.....	239
1.3	PNUMA .....	244
1.4	ONUDI .....	249
1.5	UNICEF.....	254
1.6	FAO.....	259
1.7	FMAM .....	264
1.8	OIOS.....	271
1.9	PMA .....	274
1.10	ONU Mujeres .....	279
	<b>Anexo 2: Descripción de puntuaciones en escala difusa por agencia .....</b>	<b>285</b>

## ÍNDICE DE TABLAS

Tabla 1.1. Categorías y criterios del uso de la evaluación .....	39
Tabla 1.2. Complementariedades entre Monitoreo y Evaluación .....	58
Tabla 1.3. El tipo de utilización y su factibilidad en Monitoreo y Evaluación .....	59
Tabla 2.1. Evaluación y capacidad evaluativa.....	70
Tabla 2.2. Objetivos en el desarrollo de capacidad evaluativa.....	95
Tabla 3.1. Estructura de la función de evaluación en los organismos de la Secretaría .	121
Tabla 4.1. Políticas de evaluación UNEG .....	143
Tabla 4.2. Indicadores y su graduación en escala difusa.....	154
Tabla 4.3. Sumario de puntuación en escala difusa (fuzzy set) .....	159
Tabla 5.1. Análisis de condiciones necesarias para el resultado ‘calidad de la evaluación’ .....	164
Tabla 5.2. Tabla de la verdad sobre el resultado ‘Calidad de la Evaluación’ .....	171
Tabla 5.3. Análisis de condiciones suficientes en ‘Calidad de la Evaluación’. .....	172
Tabla 5.4. Análisis de condiciones necesarias para el resultado ‘uso de la evaluación’. 183	
Tabla 5.5. Tabla de la verdad sobre el resultado ‘Uso de la Evaluación’. .....	187
Tabla 5.6. Análisis de condiciones suficientes en ‘Uso de la Evaluación’. .....	187

## ÍNDICE DE FIGURAS

Figura 1.1. Contribución del proceso de evaluación participativa a la utilización .....	34
Figura 1.2. Expansión del concepto de uso de la evaluación.....	43
Figura 1.3. Teoría integrada de influencia.....	44
Figura 1.4. Mecanismos a través de los cuales la evaluación produce influencias .....	46
Figura 2.1. Triángulo de Principios de la Declaración de París.....	67
Figura 2.2. Dimensiones de los enfoques participativos .....	73
Figura 2.3. Modelo teórico de la evaluación empoderativa .....	87
Figura 2.4. Ciclo de evaluación colaborativa anual .....	90

Figura 2.5. Modelo Multidisciplinario .....	100
Figura 2.6. Estrategias en el enfoque de inmersión colaborativa .....	101
Figura 2.7. Modelo Integrador de Capacidad Evaluativa.....	107

## ÍNDICE DE GRÁFICOS

Gráfico 5.1. Políticas/Prácticas Evaluativas y Calidad de la Evaluación .....	165
Gráfico 5.2. Institucionalización y Calidad de la Evaluación.....	167
Gráfico 5.3. Capacidad de Aprendizaje y Calidad de la Evaluación.....	168
Gráfico 5.4. La calidad de la evaluación desde el aprendizaje y las políticas/prácticas evaluativas .....	173
Gráfico 5.5. La calidad de la evaluación desde la institucionalización y las políticas/prácticas evaluativas.....	173
Gráfico 5.6. Capacidad de Aprendizaje y Uso de la Evaluación .....	184
Gráfico 5.7. Políticas/Prácticas evaluativas y Uso de la Evaluación .....	185
Gráfico 5.8. Institucionalización y Uso de la Evaluación .....	186
Gráfico 5.9. El uso de la evaluación desde la capacidad de aprendizaje y las políticas/prácticas evaluativas.....	189
Gráfico 5.10. El uso de la evaluación desde la capacidad de aprendizaje y en ausencia de institucionalización .....	189

## **Glosario**

ACNUDH: Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos

ACNUR: El Alto Comisionado de las Naciones Unidas para los Refugiados

AG: Asamblea General

AOD: Ayuda Oficial al Desarrollo

CTBTO: Comisión Preparatoria de la Organización del Tratado de prohibición completa de los ensayos nucleares

DAAT: Departamento de Apoyo a las Actividades sobre el Terreno

DAC: Development Assistant Committee

DAES: Departamento de Asuntos Económicos y Sociales

DAGGC: Departamento de la Asamblea General y de Gestión de Conferencias

DAP: Departamento de Asuntos Políticos

DFID: Departamento para el Desarrollo Internacional del Reino Unido

DG: Departamento de Gestión

DIP: Departamento de Información Pública

DOMP: Departamento de Operaciones de Mantenimiento de la Paz

DS: Departamento de Seguridad

FAO: Organización de las Naciones Unidas para la Agricultura y la Alimentación

FIDA: Fondo Internacional de Desarrollo Agrícola

FMAM: Fondo Mundial de Medio Ambiente

FMI: Fondo Monetario Internacional

FNUAP: El Fondo de Población de las Naciones Unidas

GAO: Government Accountability Office

GATE: Global Accountability and Tracking of Evaluation System

GERAAS: Global Evaluation Reports Assessment and Analysis System

GEROS: Global Evaluation Reports Oversight System

GIZ: Agencia Alemana de Cooperación Técnica

JICA: Agencia de Cooperación Internacional de Japón

JIU: Joint Inspection Unit

MAR: Management Action Record

MSIP: Plan Estratégico e Institucional de Medio Término

NORAD: Agencia Noruega de Cooperación al Desarrollo

NU: Sistema de Naciones Unidas

OACI: Organización de Aviación Civil Internacional

OAD: Oficina de Asuntos de Desarme de las Naciones Unidas

OAJ: Oficina de Asuntos Jurídicos

OCAH: Oficina de Coordinación de Asuntos Humanitarios

ODS: Objetivos de Desarrollo Sostenible

OECD: Organisation for Economic Co-operation and Development

OIEA: Organismo Internacional de la Energía Atómica

OIOS: Office of Internal Oversight Services

OIT: Organización Internacional del Trabajo

OMC: Organización Mundial del Comercio

OMI: Organización Marítima Internacional

OMM: Organización Meteorológica Mundial

OMPI: Organización Mundial de la Propiedad Intelectual

OMS: Organización Mundial de la Salud

OMT: Organización Mundial del Turismo

ONU Habitat: El Programa de Naciones Unidas para los Asentamientos Humanos

ONUDI: Organización de las Naciones Unidas para el Desarrollo Industrial

OOPS: El Organismo de Obras Publicas y Socorro de las Naciones Unidas para los Refugiados de Palestina en el Cercano Oriente

OPAQ: Organización para la prohibición de las Armas Químicas

PMA: El Programa Mundial de Alimentos

PNUD: El Programa de las Naciones Unidas para el Desarrollo

PNUMA: El Programa de Naciones Unidas para el Medioambiente

PPE: Evaluación participativa práctica

QCA: Análisis Comparativo Cualitativo

TPE: Evaluación participativa transformativa

UIT: Unión Internacional de Telecomunicaciones

UNCTAD: La Conferencia de las Naciones Unidas sobre Comercio y Desarrollo

UNDESA: Departamento de Asuntos Sociales y Económicos de las Naciones Unidas

UNEG: Grupo de Evaluación de las Naciones Unidas

UNESCO: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura

UNICEF: El Fondo de las Naciones Unidas para la Infancia

UPU: Unión Postal Universal



## Introducción

### *Justificación*

Hoy en día, la evaluación es una práctica común dentro del Sistema de las Naciones Unidas (NU). En muchos de sus organismos afines, existen unidades organizativas dotadas de recursos y reguladas por procedimientos específicos para la realización de evaluaciones. El Grupo de Evaluación de las Naciones Unidas (UNEG, por sus siglas en inglés) es un marco de referencia para estas instituciones a la hora de aplicar estándares y procedimientos relativos a la práctica evaluativa así como también un importante foro para el intercambio de experiencias entre las diferentes agencias sobre aspectos evaluativos.

A pesar de que la práctica evaluativa es una disciplina asentada en las instituciones internacionales y goza de un buen momento con respecto a etapas anteriores, también es verdad que todavía la evaluación se enfrenta a algunos desafíos importantes. Existe una percepción bastante generalizada entre la comunidad evaluadora y los equipos responsables de la gestión de los programas evaluados de que las evaluaciones tienen una utilidad limitada y que, con frecuencia, sus resultados se convierten en documentos aislados carentes de una repercusión significativa en la marcha operativa de la organización en cuestión.

En este contexto, el concepto de capacidad evaluativa emerge con fuerza en los últimos años en el entorno de la comunidad evaluadora como parte de una estrategia orientada a mejorar la calidad y facilitar el uso de las evaluaciones. Así, el marco teórico desarrollado hasta la fecha sobre capacidad evaluativa se ha centrado principalmente en dar respuesta a cuatro cuestiones: 1) ¿Qué es capacidad evaluativa?; 2) ¿Qué podemos hacer para que ocurra?; 3) ¿Cómo podemos saber si está ocurriendo? y 4) ¿Cuál es su impacto?

Muchos teóricos y teóricas de la evaluación han destinado importantes esfuerzos a especificar el significado más apropiado del concepto de capacidad evaluativa (i.e Baizerman et al. 2002; Compton et al. 2002; García-Iriarte et al. 2011; Huffman et al. 2008; King 2002; Labin et al. 2012; Stockdill, Baizerman y Compton 2002). Su punto de referencia se remonta a los denominados enfoques participativos de la evaluación (también denominados enfoques de evaluación participativa, empoderativa o colaborativa), los cuales conciben la evaluación como un proceso incluyente de responsabilidades compartidas entre las partes implicadas. Así, la concepción actual de la capacidad evaluativa se sustenta, en buena medida, sobre la base de la participación como aspecto fundamental.

A su vez, la literatura existente en la materia ha aportado al concepto de capacidad evaluativa una serie de modelos teóricos que, algunos de ellos respaldados con evidencias empíricas, contribuyen a esclarecer otros aspectos relevantes – más allá del carácter participativo previamente comentado – que completarían el concepto de capacidad evaluativa (i.e Connolly y York 2002; Cousins et al. 2004; Nielsen et al. 2011, Preskill y Boyle 2008). A pesar de que

la terminología utilizada en los modelos sobre capacidad evaluativa puede presentar ciertas diferencias, lo cierto es que hay tres factores fundamentales, también denominados condiciones, que se consideran comunes en dichos modelos: 1) institucionalización de la función de evaluación; 2) marco normativo y 3) capacidad de aprendizaje.

En el marco de las NU, la Asamblea General (AG) también se ha sumado a los enfoques teóricos que apuntan a fortalecer la capacidad evaluativa en los países e instituciones (A/RES/62/208 de 2007, y más recientemente A/RES/69/237 de 2015), como instrumento de mejora en la sostenibilidad, uso y calidad de la evaluación. A su vez, algunos de los esfuerzos internacionales más reconocidos en este sentido son las series de foros globales de alto nivel sobre la eficiencia de la ayuda organizados por la OCDE en Roma, París, Accra y Busan en los años 2003, 2005, 2008 y 2011 respectivamente. En particular la Declaración de París y el Programa de Acción de Accra han sido dos instrumentos muy relevantes para justificar la necesidad de desarrollar la capacidad evaluativa como herramienta de mejora de la efectividad de la ayuda. Dichos instrumentos abogan así con contundencia por fortalecer las capacidades de planificación, monitoreo y evaluación de las intervenciones de cooperación internacional.

En este contexto, la UNEG ha dedicado importantes esfuerzos para el desarrollo de las funciones de evaluación y la adopción armonizada de normas y estándares para la práctica evaluativa en las NU. Estas normas y estándares han sido la base para el desarrollo de varias prácticas hoy asentadas en el ámbito de la evaluación, como por ejemplo el proceso de revisión por pares (*peer review process*) que en base a tres pilares básicos (independencia, utilidad y credibilidad) otorga un papel primordial a los organismos del Sistema de NU en los procesos de valoración de la función de evaluación de los propios organismos, valorando así cada agencia del Sistema, la actuación de sus agencias hermanas (UNEG/REF/2011/1).

A pesar de la importancia otorgada en el ámbito internacional al fortalecimiento de la capacidad evaluativa y, en particular, a los esfuerzos desplegados por la UNEG para el desarrollo de funciones de evaluación sólidas, los efectos que han generado dichas iniciativas son todavía insuficientes (JIU/REP/2014/6).

Aunque en la actualidad existe una amplia literatura especializada en relación al concepto de capacidad evaluativa así como modelos teóricos que tratan de explicar las causas y efectos de su presencia o ausencia, han sido pocos los estudios donde se analizan casos concretos y experiencias prácticas. En consecuencia, todavía no existe una sólida evidencia empírica que permita desvelar hasta qué punto los esfuerzos para fortalecer la capacidad evaluativa en las instituciones han llevado al cumplimiento de sus objetivos, que en última instancia se representan por la mejora en la calidad de las evaluaciones y el uso rutinario de la evaluación en la toma de decisiones. En contraste, los informes existentes publicados por los órganos de control de las NU (i.e., A/70/72), aluden a la deficiente capacidad evaluativa de sus instituciones y al limitado poder de influencia de la práctica evaluativa en la toma de decisiones.

De lo anterior subyacen algunos interrogantes que han motivado este trabajo: ¿Porqué tras tantos esfuerzos orientados a reforzar la capacidad evaluativa como mecanismo de cambio y mejora dentro de las organizaciones, los resultados obtenidos no son satisfactorios? ¿Cómo se explica la percepción generalizada de que la práctica evaluativa aún no ha conseguido influenciar de manera significativa el proceso de toma de decisiones estratégicas dentro de las diferentes instituciones? ¿Cuáles son los elementos que se han de reforzar para que la evaluación se convierta en una práctica sostenible que favorezca el aprendizaje organizacional?

### *Objetivos*

Esta tesis tiene como objetivo central contribuir a un mayor conocimiento de los factores facilitadores de la capacidad evaluativa (de acuerdo al marco teórico existente) para lograr una mayor utilización de la evaluación y mejorar su calidad dentro de las instituciones del Sistema de NU. En concreto, los objetivos específicos de la tesis están encaminados hacia:

1. Contribuir a un mayor conocimiento sobre cuáles son las condiciones más relevantes y necesarias para favorecer el uso y la calidad de las evaluaciones realizadas en el marco de organizaciones multidisciplinares como las de las NU, en contraste con las condiciones generalmente aceptadas por el marco teórico desarrollado en esta área.
2. Aportar recomendaciones para el fortalecimiento del marco teórico existente sobre capacidad evaluativa, agregándole elementos que ponderen la importancia relativa de las diferentes condiciones orientadas a facilitar la calidad y uso sostenible de las evaluaciones.

### *Hipótesis*

La hipótesis que se contrasta en este trabajo se centra en la idea de que la cultura de aprendizaje organizacional es un elemento clave dentro de las instituciones para favorecer la utilización de la evaluación. Sin embargo, la cultura de aprendizaje organizacional por si sola puede no ser suficiente para realizar una práctica evaluativa de calidad y sostenible en las instituciones. Esta cultura de aprendizaje debe ir acompañada de un sólido marco normativo en materia de evaluación, para convertirse en un factor fundamental para el logro de la capacidad evaluativa que se necesita en las instituciones de las NU, es decir, una capacidad evaluativa que permita la mejora de la calidad de las evaluaciones y su influencia en la toma de decisiones. Hasta ahora, esta combinación de factores que resulta de la cultura de aprendizaje organizacional junto con el marco normativo, no ha sido reconocida por los modelos teóricos que analizan la capacidad evaluativa como un elemento de crucial importancia.

Así, la hipótesis central a contrastar es la siguiente: la capacidad de aprendizaje se constituye como condición necesaria para lograr resultados satisfactorios de utilización de la evaluación, y si ésta es combinada con un

marco normativo adecuado, constituye entonces una fórmula efectiva para producir evaluaciones de calidad en las organizaciones.

Sin cultura de aprendizaje organizacional, el marco normativo y la institucionalidad de la función de evaluación son factores que pierden efectividad y difícilmente van a lograr por sí mismos la plena utilización de la función de evaluación. Así, deficiencias en la cultura de aprendizaje organizacional producen efectos negativos en el logro de los resultados que persigue la capacidad evaluativa, a pesar de que las otras condiciones del marco teórico estén presentes (especialmente las relativas a fortalecer la institucionalización de la función de evaluación). La cultura de aprendizaje organizacional es, por tanto, un eje central en el desarrollo pleno de la capacidad evaluativa y su presencia facilita que los demás factores sean efectivos.

En consecuencia y dado que la práctica evaluativa, aunque presente en las instituciones, no está todavía asentada como herramienta de aprendizaje organizacional, la capacidad evaluativa de las organizaciones es aún débil, hecho que limita la influencia de las evaluaciones en la toma de decisiones.

Aplicada al marco de NU, las sub-hipótesis a contrastar son:

1. La evaluación en NU ha experimentado grandes avances en cuanto a su consolidación como disciplina y en cuanto a la dotación de recursos para su establecimiento organizativo y programático. Se han fortalecido las unidades de evaluación como estructuras funcionalmente independientes con el fin de proporcionar evaluaciones imparciales y de calidad. Otros aspectos como el de la integración de la evaluación como parte del aprendizaje organizacional están aún pendientes. Debido a ello, las evaluaciones siguen teniendo un carácter predominantemente aislado y este hecho limita su utilidad y su influencia.
2. Entre los diferentes enfoques aplicados por las organizaciones de NU se puede distinguir entre aquellos que han priorizado el fortalecimiento de la función de evaluación desde su institucionalidad con la rendición de cuentas como propósito principal, frente a los que han primado los enfoques consultivos, participativos y donde la diseminación de la evaluación tiene un claro componente de aprendizaje. El segundo grupo ofrece mejores resultados en cuanto a sostenibilidad ya que la institucionalidad de la función de evaluación no lleva necesariamente a fortalecer la utilización de la misma.
3. Los organismos con una mayor capacidad de aprendizaje organizacional y que han combinado este factor con el establecimiento de un marco normativo sólido muestran una mayor capacidad evaluativa. Existe evidencia de que en dicho contexto, la evaluación de calidad forma parte de la toma de decisiones en materia programática (programas evaluados) así como en relación a las políticas institucionales.

## Metodología

Para poder desarrollar las hipótesis planteadas, se realiza en primer lugar una revisión de la literatura especializada sobre las teorías y modelos existentes en relación al uso de la evaluación y la capacidad evaluativa de las organizaciones.

En segundo lugar, para contrastar las hipótesis planteadas en base a evidencias empíricas, se desarrolla un análisis comparativo sobre una muestra representativa de agencias del sistema de las NU. Este análisis se centra en la capacidad evaluativa de las agencias y trata de poner de manifiesto que aquellos organismos que gozan de una mayor capacidad de aprendizaje organizacional, si ésta viene acompañada de un sólido marco normativo, el resultado que la combinación de ambos genera es la producción de evaluaciones de calidad que proporcionan hallazgos más fáciles de absorber por las organizaciones y, por tanto, capaces de generar los cambios estructurales y/u operativos correspondientes que ameriten en cada caso.

La muestra seleccionada para llevar a cabo este análisis comparativo está compuesta por diez agencias del Sistema de NU las cuales gozan todas ellas de una importante tradición en relación al desarrollo e implementación de la práctica evaluativa y que, a su vez, han estado sometidas a un proceso de revisión por pares de sus funciones de evaluación<sup>1</sup>. A su vez, el conjunto de estas diez agencias es representativo de las diferentes formas en que la evaluación se desarrolla en el Sistema de NU.

El método seleccionado para realizar la comparación se denomina Análisis Comparativo Cualitativo (*Qualitative Comparative Analysis, QCA*). Este método ha sido seleccionado por ofrecer una serie de características que se adaptan de forma óptima a las exigencias de este trabajo. QCA es un método comparativo configuracional (*Configurational Comparative Method*) desarrollado por el sociólogo estadounidense Charles Ragin en la década de los 80 que permite establecer comparaciones entre entidades o casos (en este caso agencias de las NU) a partir de la transformación de las mismas en diferentes configuraciones o grupos de variables, componentes o - por ser fieles a la terminología empleada - condiciones, que agrupadas de una u otra manera, contribuyen a generar un resultado determinado de interés para la investigación en cuestión.

En el caso que nos atañe, las condiciones vienen representadas por los factores que componen la capacidad evaluativa (identificadas y desarrolladas en el marco teórico) y los resultados de interés son el uso final de las evaluaciones y su calidad. Cada agencia, en nuestra aplicación, despliega una calificación determinada en relación tanto a las condiciones presentes como al resultado. De esta forma, podremos comparar, por ejemplo, agencias que muestren un resultado positivo (un gran uso de las evaluaciones que producen, una gran calidad) a partir de una combinación determinada de condiciones, con otras agencias cuyo resultado no sea tan alentador (sus evaluaciones no

---

<sup>1</sup> Peer review process

influyen tanto en la organización) y que despliegan una combinación diferente a la anterior en cuanto a las condiciones presentes.

Con un sustento empírico sólido, QCA permite comparar todos los casos (agencias) desvelando las combinaciones de condiciones que son necesarias y/o suficientes para la obtención del resultado de interés (uso de las evaluaciones y calidad de las evaluaciones). Una condición será considerada necesaria si ésta se encuentra siempre presente cuando ocurre el resultado. Ésta será suficiente si, en contraste con lo anterior, siempre que la condición está presente, entonces se produce el resultado (aunque puede ocurrir que produciéndose el resultado, la condición no esté presente).

Este tipo de análisis nos permite entender hasta qué punto una condición juega un rol determinante en la consecución de un resultado particular. En el caso que nos concierne, el marco teórico explicativo de la capacidad evaluativa, como se mostrará, no ha ponderado hasta qué punto los factores (que conforman la capacidad evaluativa bien de forma individual o en combinación con otros factores) son necesarios y/o suficientes para que las evaluaciones tengan calidad y sean útiles e influyentes en la toma de decisiones de las organizaciones.

A partir de los resultados obtenidos con la aplicación de la técnica QCA, se elaborará un análisis detallado de las condiciones más importantes para la consecución de los resultados, así como de las agencias que, por los resultados y combinación de condiciones que ofrecen, sean más interesantes desde un punto de vista analítico.

### *Estructura*

Para confirmar las hipótesis planteadas, se analizará detalladamente la relación existente entre los factores facilitadores para el desarrollo de la capacidad evaluativa en el marco de NU y su grado de influencia en el uso y calidad de las evaluaciones.

Así, este trabajo se inicia con un primer capítulo donde se analiza el marco teórico existente sobre evaluación y su utilidad. En concreto, se examinará el concepto de utilización en evaluación desde su evolución histórica, los diferentes tipos de utilización que existen y cómo los debates conceptuales actuales han evolucionado hacia un concepto más multidimensional de la evaluación, donde ésta es parte integrante de un sistema de información que nutre de forma sistemática la toma de decisiones.

El segundo capítulo se centra en el concepto de capacidad evaluativa, revisando en detalle el marco teórico en el que se ubica así como las definiciones existentes de dicho concepto y sus principales propósitos. Concretamente, se indagará sobre cuáles son las relaciones existentes entre evaluación, conocimiento y capacidad evaluativa. A su vez, se procederá a la descripción de los enfoques participativos en el marco teórico y su influencia en la teoría existente así como los modelos conceptuales más relevantes de capacidad evaluativa.

El tercer capítulo describe cómo se desarrolla la función de evaluación en el Sistema de NU, cuáles son sus principios rectores, su marco normativo y, en particular, los mecanismos existentes para el fortalecimiento de la capacidad evaluativa y en consecuencia de las funciones de evaluación.

En correspondencia con las hipótesis a contrastar, en el cuarto capítulo se desarrolla una propuesta de marco metodológico para el análisis de la capacidad evaluativa en el Sistema de NU. Para ello, se recogen los principales factores identificados en el marco teórico como relevantes para explicar resultados positivos en las estrategias de capacidad evaluativa. Dichos factores, teniendo en cuenta que nos encontramos en el ámbito de aplicación del Sistema de NU, serán alineados consecuentemente con el marco de normas y estándares UNEG.

A continuación se procederá a la descripción del método de análisis comparativo cualitativo y, en concreto, se detallarán las razones que justifican la elección de este método en la investigación de la presente tesis. Además se desarrollará paso por paso el procedimiento de puntuación en escala difusa en base a los indicadores seleccionados para realizar el análisis empírico de las condiciones y resultados vinculados al marco teórico de la capacidad evaluativa. Sin embargo, la descripción detallada de las agencias objeto de estudio en función de su mandato, gobernabilidad, marco estratégico, recursos, función de evaluación y proceso de revisión por pares se ha trasladado al Anexo 2 para facilitar la lectura y comprensión del documento.

El capítulo quinto analiza los resultados obtenidos con el método de análisis comparativo cualitativo y, en base a estos resultados, se discute y se contrasta la capacidad evaluativa de las instituciones objeto de análisis en relación a los factores determinantes propuestos en el marco metodológico, examinando sus efectos sobre el grado de influencia en el uso y la calidad de la evaluación. La información se analiza en base a la identificación de relaciones causales para las condiciones necesarias y suficientes.

El último capítulo (capítulo sexto) desarrolla las conclusiones de la presente investigación y profundiza en los factores que fortalecen la capacidad evaluativa y, por tanto, facilitan la calidad y el uso de la evaluación como práctica sostenible capaz de influir en la toma de decisiones organizacionales.

Por último, cabe mencionar que a lo largo de la tesis se ha hecho un uso no sexista del lenguaje intentando dar presencia y voz al trabajo de investigación y análisis empírico realizado por las mujeres evaluadoras, muy presentes y activas en la literatura y en la práctica evaluativa.



## 1. El valor de la evaluación desde su utilidad

### 1.1. La evaluación como concepto general

A pesar de que el pensamiento evaluativo se puede considerar un aspecto inherente a nuestra condición humana, la evaluación formal y sistemática - la que se realiza desde la práctica profesional - es un fenómeno reciente que tiene sus inicios en Estados Unidos en el siglo XX. Los campos de la educación y la salud fueron los primeros en incorporar estudios sistemáticos de evaluación para valorar los resultados obtenidos en dichas áreas. Así, investigadoras e investigadores como Weiss (1988a), Rossi et al. (2004) y Patton (2008) han identificado y analizado estudios de evaluación publicados con anterioridad a 1950, principalmente provenientes de fundaciones privadas de carácter social e instituciones ligadas a la salud y a la educación.

Tal y como algunos manuales de evaluación apuntan, no existe una definición comúnmente aceptada de evaluación (Shadish 1994) y las numerosas definiciones existentes, reflejan no solo la variedad de aspectos inherentes a la evaluación que son destacados como primordiales, como su valor, utilidad o los métodos utilizados (Alkin 2004), sino también la enorme diversidad de criterios en cuanto a sus fundamentos teóricos (Sechrest 1994; Worthen 1999; Scriven 1994a) y sus resultados esperados (Scriven 1994b; Stufflebeam 1994).

Las perspectivas y definiciones sobre evaluación son, por tanto, numerosas y han sido tema de referencia y debate en numerosos manuales de evaluación (Vedung 1997; Clarke 1999; Rossi et al. 1985; Rossi et al. 2004). Uno de los aspectos que recogen la mayoría de las definiciones es precisamente la condición de juzgar el valor o el mérito del objeto evaluado (Rist y Morras 2004). En palabras de Rist y Morras (2004, pág. 8), es *“precisamente éste aspecto valorativo el que se usa para diferenciar evaluación de otras formas de investigación y monitoreo”*.

En un intento de armonizar las diferentes definiciones en el ámbito de la cooperación y la ayuda al desarrollo, el Comité de Asistencia al Desarrollo (*Development Assistant Committee, DAC*) de la OECD define el término evaluación como:

*“la apreciación sistemática de un proyecto, programa o política en curso o concluido, de su diseño, su puesta en práctica y sus resultados. El objetivo es determinar la pertinencia y el logro de los objetivos, así como la eficiencia, la eficacia, el impacto y la sostenibilidad para el desarrollo. Una evaluación deberá proporcionar información creíble y útil, que permita incorporar las enseñanzas aprendidas en el proceso de toma de decisiones de beneficiarios y donantes”* (OECD 2000, pág. 21)

Esta definición facilita el análisis de cuatro elementos clave de la evaluación desde un punto de vista conceptual. Los elementos son: 1) el objeto evaluado, 2) el foco de análisis, 3) los principios, estándares y criterios evaluativos, y 4) el propósito de la evaluación. A continuación, se procede a una breve descripción

de los elementos anteriores, con el fin de profundizar más en el significado de la evaluación como concepto.

## 1. El objeto evaluado.

Las evaluaciones pueden ser clasificadas en función del objeto evaluado. Se puede diferenciar así entre las evaluaciones de políticas públicas y las evaluaciones de programas o proyectos, las cuales de acuerdo con Patton (1997) se corresponden con las intervenciones más usuales y relevantes.

Evaluaciones de políticas públicas: Las políticas públicas en general recogen el conjunto de objetivos, decisiones y acciones que lleva a cabo un gobierno con el fin último de solucionar los problemas que en un momento determinado tanto la ciudadanía como el propio gobierno consideran prioritarios (Tamayo Sáez 1997). Así, las políticas públicas se cristalizan en programas y proyectos a los que se les asignan recursos para su puesta en práctica. Mientras que la política pública se desenvuelve en un plano más general a nivel político, los programas se asientan en planes específicos de acción con mayor o menor capacidad de dar solución a problemas o necesidades determinadas (CIPPEC 2012).

Evaluaciones de programas y proyectos: los proyectos se refieren a intervenciones individuales que persiguen alcanzar objetivos específicos a partir de una serie de recursos definidos dentro de un calendario de ejecución (OECD 2000). Son considerados como la unidad mínima de asignación de recursos para el logro de unos objetivos específicos y agrupados, constituyen la categoría de programas. Los proyectos deben tener una población objetivo definida en función de la necesidad que se pretende satisfacer, una localización espacial y tiempos de inicio y finalización predefinidos (Coen y Franco 2005).

La evaluación de políticas públicas es una práctica habitualmente vinculada al ámbito gubernamental dentro de la esfera estatal<sup>2</sup> mientras que las evaluaciones de programas y proyectos se encuentran más vinculadas a esferas de carácter supranacional. Desde la emergencia de la amplia red institucional a nivel internacional después de la Segunda Guerra Mundial, las evaluaciones de programas y proyectos se han consolidado como una práctica mayoritariamente circunscrita al ámbito de la cooperación internacional al desarrollo, tal y como se analiza en el capítulo 2 (apartado 2.2).

## 2. El foco de atención.

En función de cual sea el foco de atención de la evaluación, bien sea su proceso o sus resultados, la clasificación más relevante en la literatura evaluadora ha sido la efectuada por Scriven en 1967. En dicha clasificación, Scriven plantea los siguientes dos tipos de evaluaciones:

---

<sup>2</sup>Aunque el término de política pública ha sido tradicionalmente vinculado al ámbito doméstico, es importante destacar que en la Unión Europea cada vez más es más frecuente encontrar que el desarrollo de políticas públicas se realiza en la esfera comunitaria, lo que conlleva a que la práctica evaluativa de políticas públicas también se lleve a cabo a nivel europeo.

Evaluación formativa: Evaluación formativa es aquella cuyo objetivo es mejorar el desempeño del objeto evaluado. Ésta, generalmente “*se realiza durante la fase de implementación de un proyecto o programa*” (OECD/DAC 2002, pág. 23). Tiene un carácter continuo, ligado a la implementación del objeto de estudio, con el fin de localizar deficiencias y desviaciones sobre los planes previamente especificados cuando aún se está a tiempo para remediarlas mediante la introducción sobre la marcha de las rectificaciones adecuadas. El principal objetivo de la evaluación formativa es la mejora del producto final y se aleja de consideraciones sancionadoras (énfasis en el factor aprendizaje). En ocasiones, también es llamada evaluación de procesos.

En este sentido, es frecuente que las evaluaciones formativas sean llevadas a cabo bien por personas de la propia organización, o bien por expertos comisionados por los órganos responsables de la dirección y gestión del programa evaluado, para poder así llevar a cabo con mayor facilidad las rectificaciones necesarias (Scriven 1991). Se puede afirmar que este tipo de evaluación representa una tarea que forma parte intrínseca del desarrollo del objeto de estudio, es decir, del desarrollo del programa o proyecto evaluado.

Evaluación sumativa: Evaluación sumativa es aquella evaluación que se realiza al final de una intervención para determinar en qué medida se produjeron los resultados previstos. “*El objetivo final de la evaluación sumativa es proporcionar información sobre el valor del programa*” (OECD/DAC 2002, pág. 35). La evaluación sumativa (también llamada evaluación de resultados) proporciona información sobre la eficacia del objeto evaluado, tratando de dilucidar si éste ha conseguido cumplir los objetivos para los cuales fue diseñado. Generalmente, la evaluación sumativa se produce una vez el programa, política o proyecto ha terminado su ejecución y, por tanto, los resultados obtenidos son verificables. Los objetivos de la evaluación sumativa van así vinculados a la rendición de cuentas y normalmente es llevada a cabo por personal externo para garantizar su independencia de las partes implicadas en el objeto evaluado. Como denota Scriven, “*la evaluación sumativa es realizada en última instancia para la toma de decisión, un ejercicio que necesita de conclusiones evaluativas más allá del desarrollo e implementación del objeto evaluado*” (Scriven 1991, pág. 20).

En consecuencia, la principal diferencia entre ambos tipos de evaluación (evaluación formativa y evaluación sumativa) reside en la dimensión temporal (momento específico en el que se desarrolla la evaluación) y en el objetivo planteado en relación a la evaluación en sí misma. Para el caso de la evaluación formativa, el foco de atención se centra en el proceso de implementación (momento en el que se realiza la misma) de cara a mejorar el producto final, mientras que la evaluación sumativa se desarrolla una vez concluido el proyecto evaluado y se centra en los resultados obtenidos de cara al establecimiento de juicios de valor sobre la intervención evaluada para la consecuente rendición de cuentas.

La validez y relevancia de esta clasificación (formativa o sumativa) no ha estado exenta de debate (Cronbach et al 1980; Scriven 1991). Mientras que Scriven es firme defensor de esta clasificación en cuanto a que “*todo lo que no*

*sea evaluación formativa, es evaluación sumativa*”, así como de la importancia de la evaluación sumativa frente a la formativa ya que para él “*la evaluación formativa no merece la pena a no ser que al menos sirva como preámbulo de una buena evaluación sumativa*” (Scriven 1991, pág. 53); Cronbach, en contraste, ha manifestado su preferencia por la evaluación formativa al considerar ésta de mayor utilidad para el éxito de los programas y proyectos evaluados así como para las instituciones que los implementan (Cronbach et al. 1980).

Otros autores, sin embargo, entre los que se encuentran Patton (1996) y Chen (1996), consideran que esta clasificación puede resultar demasiado rígida e incluso incompleta, y que ambas categorías (sumativa y formativa) no son necesariamente excluyentes. En este debate, Rossi et al. (2004) introducen un tipo adicional de evaluación, la Evaluación Prospectiva, la cual tiene como objetivo valorar la posibilidad de obtener resultados positivos de proyectos, programas o políticas que todavía no están en ejecución. Es similar a lo que podría considerarse como ‘exámenes de evaluabilidad’. Este tipo de evaluaciones se sustentan en los hallazgos procedentes de evaluaciones anteriores, los cuales aportan elementos importantes para estimar en qué medida se podrán obtener resultados positivos del proyecto o programa evaluado. En este sentido, es frecuente que este tipo de evaluaciones incluyan una reconstrucción del proceso de elaboración del programa evaluado y recopilen resultados procedentes de experiencias similares implementadas en contextos afines.

### 3. Evaluación de evaluaciones (meta-evaluación).

Independientemente de las clasificaciones anteriores, la práctica evaluativa se ha expandido en relación a su unidad de análisis. Así, más allá de estudiar políticas, programas y proyectos, una línea de acción dentro de este campo de las evaluaciones, se concentra en evaluar las propias evaluaciones en sí mismas y, en definitiva, la propia práctica evaluativa. Son las denominadas meta-evaluaciones.

Es difícil afirmar con certeza dónde y cuándo se origina el término de ‘meta-evaluación’. Scriven en 1967 y posteriormente en 1991 define meta-evaluación como la evaluación de evaluaciones e indirectamente como la evaluación de evaluadores. De acuerdo a Scriven, “*las meta-evaluaciones representan una obligación científica e incluso ética, debido a la trascendencia que la práctica evaluativa puede llegar a tener en el funcionamiento de instituciones clave en nuestra sociedad, las cuales, en última instancia, producen resultados que afectan al bienestar de las personas*” (1991, pág. 228). La meta-evaluación es comúnmente considerada como una investigación sistemática cuyo objetivo es emitir un juicio acerca de la calidad y de los méritos de las evaluaciones.

Entre las primeras publicaciones sobre meta-evaluaciones se destaca la serie de artículos editados por Niemiec y Walberg en 1992. Estos analizaban un total de trece evaluaciones de programas educativos realizadas entre 1975 y 1987 en los que se había introducido herramientas informáticas (*computer assisted instruction*) como parte del método de aprendizaje. Los artículos incluyeron un

análisis de las fortalezas y debilidades de los procesos evaluativos desarrollados en dichos ámbitos educativos con el fin de mejorar la práctica evaluativa en sí misma. Posteriormente, la Revista Americana de Evaluación (*American Journal Evaluation*) introdujo en 1999 una sección llamada 'Evaluando Evaluaciones' (*evaluating evaluations*) que pretendía incorporar artículos relacionados con meta-evaluaciones en diversas áreas de interés. Tal y como se describe en la introducción de la sección, se trataba de mejorar la práctica de la evaluación a la vez que demostrar la utilidad de las meta-evaluaciones.

Existe cierta confusión entre la definición de meta-evaluación y síntesis de evaluaciones. Algunos estudios se autodenominan meta-evaluaciones cuando en realidad son más analíticos evaluativos (Ashworth et al. 2004). Para diferenciarlos, Stufflebeam y Shinkfield (2007) puntualizan que las meta-evaluaciones analizan el mérito y el valor de las diferentes evaluaciones, mientras que la síntesis de evaluaciones es más un resumen de un grupo de estudios de evaluación que poseen un patrón común en algún punto del proceso de investigación.

Profundizando en esta línea, Stufflebeam y Shinkfield en 2007 distinguen dos tipos de meta-evaluaciones: las meta-evaluaciones proactivas, las cuales se realizan por expertos externos a un proceso determinado de evaluación para ayudar al equipo evaluador a realizar de forma más informada y contrastada su tarea de diseñar e implementar la evaluación asignada. Esta práctica se alinea con las evaluaciones formativas anteriormente mencionadas, ya que los insumos de la meta-evaluación serán empleados por los propios evaluadores para mejorar el proceso evaluativo en el cual se encuentran inmersos. En contraste, las meta-evaluaciones retroactivas ofrecen resultados a una audiencia diferente. Son aquellas que sirven principalmente para facilitar a los clientes o en general a los receptores de la meta-evaluación (por ejemplo donantes u órganos de gobierno en el caso de las NU), la valoración de la práctica evaluativa (representada por un conjunto determinado de evaluaciones), una vez que las evaluaciones correspondientes han concluido. Este último tipo es análogo a las evaluaciones sumativas.

Independientemente del tipo de meta-evaluación, el fin último de éstas consiste en mejorar la práctica evaluativa. Stufflebeam, en este sentido afirmó en 1987 que *"el objetivo de la meta-evaluación es asegurar la calidad de los servicios evaluativos, evitar prácticas ilegales o que no son de interés público, señalar el camino para el perfeccionamiento de la profesión de evaluación y promover una mayor comprensión de la práctica evaluativa"* (1987, pág. 52).

#### 4. Códigos normativos, principios y criterios evaluativos

Para discernir el concepto general de evaluación hay que observar con detenimiento los códigos normativos, principios y criterios evaluativos que rigen la práctica evaluativa y que forman parte intrínseca de la misma.

El término código hace referencia al conjunto de reglas o preceptos que rigen el funcionamiento de cualquier institución. En el campo de la evaluación se

mantiene esta idea y a nivel orientativo dentro de esta categoría se incluyen: i) los principios (*principles*), los cuales son consideraciones de tipo ético y moral que manifiestan cómo deben comportarse los evaluadores; ii) los estándares (*standards*), que están más enfocados a los requisitos que debe cumplir una evaluación de calidad; y iii) las orientaciones (*guidelines*), que son un conjunto de disposiciones que exponen la mejor manera de cumplir con los estándares y con los principios (Bustelo 2004).

En la práctica evaluativa, la división de la categoría de códigos explicada anteriormente (en principios, estándares y orientaciones) no siempre es adoptada por las organizaciones y, con cierta frecuencia, la línea divisoria entre las diferentes subcategorías es difusa. Aquello que para una organización es considerado un principio, en otra, puede haber formar parte de sus estándares y viceversa, lo que contribuye a generar cierta confusión en este campo. En ocasiones, también es difícil diferenciar entre principios y orientaciones dentro de un mismo organismo.

Generalmente, los códigos normativos tienen una misma audiencia unida por su especialidad profesional (por ejemplo, el código ético en periodismo o abogacía). Este es uno de los puntos donde un código se diferencia de una política. Mientras que los códigos normativos dirigen sus preceptos, generalmente de trasfondo ético, a una audiencia que puede cubrir diferentes instituciones que desempeñan funciones en una misma área, la política va generalmente destinada a individuos pertenecientes a una misma institución. Así, los principios, estándares y orientaciones en ocasiones se convierten en la fuente de la que se nutren o derivan las políticas institucionales.

Tal y como describe Bustelo en 2004, la primera función de los códigos normativos en el ejercicio de una actividad profesional, también conocidos como códigos deontológicos, consiste en introducir un componente ético de conducta dentro de la actividad profesional en cuestión. La segunda función se centra en la creación de un sentido de pertenencia e identificación a una determinada comunidad profesional.

El surgimiento de códigos de evaluación ha estado intrínsecamente relacionado con un mayor grado de institucionalización y consolidación profesional de la función de evaluación (Ballart 2003). A la cabeza en el desarrollo de códigos en este ámbito, se encuentran países como Estados Unidos, Canadá, Noruega, Australia, Japón y el Reino Unido, los cuales se caracterizan por ser países con amplia experiencia en la evaluación de políticas educativas y sociales y, a su vez, por ser importantes donantes en el ámbito de la cooperación para el desarrollo, tanto a nivel nacional como a través de su participación en organizaciones multilaterales.

El primer código normativo en el área de evaluación de alcance nacional tuvo lugar en los Estados Unidos en 1981 y fue establecido por el Comité Conjunto de Estándares para la Evaluación de la Educación (*Joint Committee on Standards for Educational Evaluation*). Este amplio código de conducta que cubría el sector educativo del país, fue posteriormente revisado en 1994 y en 2011 y actualmente se encuentra en vigor. Por su parte, la Asociación

Americana de Evaluación (1994, revisada en 2004)<sup>3</sup>, la Asociación Canadiense de Evaluación<sup>4</sup>(1996), la Asociación Australiana de Evaluación<sup>5</sup> (1997), y otras asociaciones en Europa como la Sociedad de Evaluación alemana (2001), francesa (2003) y británica (2004)<sup>6</sup> lanzaron sus propios códigos normativos.

A nivel internacional, los Principios para la Evaluación de la Cooperación al Desarrollo (1991)<sup>7</sup>del Comité de Asistencia Técnica (Development Assistant Committee) de la OECD (OECD/DAC) han guiado las evaluaciones impulsadas por numerosas instituciones de asistencia para el desarrollo, entre las que se incluyen agencias de cooperación gubernamental como el Departamento para el Desarrollo Internacional del Reino Unido (DFID), la Agencia Alemana de Cooperación Técnica (GTZ), la Agencia Noruega de Cooperación al Desarrollo (NORAD) y la Agencia de Cooperación Internacional de Japón (JICA) entre otras. A su vez, dichos principios han servido como base para la formulación de las normas y estándares de otras organizaciones multilaterales como las del Sistema de Naciones Unidas en 2005, desarrolladas por la UNEG.

Además de los códigos y principios normativos, en un análisis exhaustivo del concepto general de evaluación cobran a su vez especial relevancia los denominados criterios de evaluación. Desarrollados por la OECD/DAC (OECD 1991, 2000), éstos representan una serie de puntos críticos a considerar para poder emitir de forma sustentada un juicio de valor sobre las intervenciones evaluadas. Así la OECD/DAC define los siguientes 5 criterios evaluativos:

**Pertinencia:** Medida en que los objetivos de una intervención para el desarrollo son congruentes con los requisitos de los beneficiarios, las necesidades del país, las prioridades globales y las políticas de los asociados y donantes.

**Eficacia:** Medida en que se lograron o se espera lograr los objetivos de la intervención para el desarrollo, tomando en cuenta su importancia relativa. También considerada como una medida del alcance de los objetivos de una actividad sin tener en cuenta los costes en los que se incurre para obtenerlos (Alonso y Mosley 1999).

**Eficiencia:** Medida en la que los recursos e insumos utilizados (fondos, plazos) se han transformado en resultados específicos. Optimizar la eficiencia contiene un matiz diferenciador con respecto al punto anterior (eficacia) al implicarla búsqueda de una combinación óptima de recursos financieros, materiales, técnicos, naturales y humanos que permitan obtener los resultados previstos (Alonso y Mosley 1999).

**Impacto:** Efectos de largo plazo positivos y negativos, primarios y secundarios, producidos directa o indirectamente por una intervención, de forma intencionada o involuntaria. La complejidad de aplicar este criterio en el desempeño de una evaluación se deriva de la amplitud del propio concepto de

---

<sup>3</sup>Guiding Principles for Evaluators, de la Asociación Americana de Evaluación.

<sup>4</sup>Guidelines for Ethical Conduct, de la Sociedad Canadiense de Evaluación.

<sup>5</sup>Guidelines on Ethical Conduct of Evaluation, de la Sociedad Australasiática de Evaluación.

<sup>6</sup><http://www.sfe.asso.fr>, <http://www.degeval.de> y <http://www.evaluation.org.uk>

<sup>7</sup>Principles for Evaluation of Development Assistance 1991 OECD

impacto y de la consideración por parte del mismo de efectos involuntarios no necesariamente fáciles de identificar. De acuerdo con Alonso y Mosley (1999, pág. 284), *“medir el impacto de una evaluación no se limita al estudio del alcance de los efectos previstos, tampoco no se circunscribe al análisis de los efectos deseados y no se reduce al estudio de dichos efectos sobre la población beneficiaria. Tiene un alcance mayor que incluye eventuales efectos imprevistos, no deseados y que afectan a otros colectivos diferentes de la población meta”*.

Sostenibilidad: Este criterio, considerado ampliamente por la mayoría de los actores involucrados en el sector de la cooperación al desarrollo, hace referencia al grado en el que los efectos positivos derivados de la intervención en cuestión, perduran una vez que se ha retirado la ayuda y la intervención en sí misma ha concluido.

## 5. El propósito de la evaluación

Definir con claridad a priori el propósito que persigue una evaluación determinada, aunque aparentemente pueda parecer algo elemental, no siempre es evidente y representa un punto crucial para a posteriori determinar el valor de la evaluación en sí misma. Patton sostiene que *“tener claro cuál es el propósito de una evaluación es algo fundamental para una posterior evaluación de la propia evaluación, al ser éste un punto de referencia crucial para nuestra propia rendición de cuentas”* (Patton 2012, pág.113)

El conocimiento del propósito por el que se hace la evaluación lleva directamente a la identificación de quién es el usuario final de la información proveniente de la evaluación. Como Alkin menciona, una de las primeras preguntas a realizarse ante el diseño de una evaluación es: *“Evaluación: ¿Quién la necesita? ¿A quién le importa?”* (Alkin 1975, título del artículo). De la misma manera, el propósito de la evaluación está relacionado con su potencial utilización: *“el propósito por el que se haga una evaluación, condiciona el uso que se haga de ella”* (Chelimsky 1997, pág.18).

La relación entre las tres variables mencionadas (propósito, usuario y uso de la evaluación) es más clara cuando se desgranar los principales propósitos comúnmente aceptados por la literatura evaluadora (Scriven 1991, Vedung 1997). La literatura recoge así cuatro posibles propósitos generales de una evaluación: i) la rendición de cuentas, ii) la mejora del objeto evaluado, iii) la generación del conocimiento y iv) motivos estratégicos. A continuación, nos centramos brevemente en cada uno de ellos:

La rendición de cuentas: Según Vedung, en las evaluaciones orientadas a la rendición de cuentas, *“la información y los juicios evaluativos se producen para tomar decisiones en torno a la continuación, expansión, reducción o cancelación de un programa”* (Vedung 1997, pág. 102).

La integración de este concepto en la práctica evaluativa tiene su origen en los años 70, cuando la limitación de recursos en tiempos de crisis económica demandaba la necesidad de aplicar en la evaluación, estándares propios del

sector privado relacionados con la valoración del rendimiento y la medición de resultados obtenidos en las instituciones públicas (Blalock 1999, Davies 1999). Pese a cierto grado de vaguedad inherente al concepto de rendición de cuentas, es bastante improbable encontrar opositores a la rendición de cuentas como un propósito legítimo de la práctica evaluativa, especialmente al movernos en entornos con recursos finitos. Como Cronbach ya identificó en 1980: *“aunque sea algo indefinido y sus consecuencias estén todavía sin examinar, la rendición de cuentas es percibida positivamente. ¿Hay alguien que se atreva a oponerse a la rendición de cuentas?”* (Cronbach et al., 1980, pág. 133)

Evaluadores y evaluadoras como Charlton 2002, Greene 1999 y Perrin 2007 han ahondado en la indefinición del concepto de rendición de cuentas y hacen una distinción entre una concepción tradicional vinculada a la democracia participativa y otra más tecnócrata. En este sentido, coinciden en que la rendición de cuentas ha ido gradualmente perdiendo su tradicional énfasis en el reconocimiento de que existe una responsabilidad común en las decisiones propias de un entorno de gobernabilidad democrática, transitando hacia una concepción más tecnocrática, donde la rendición de cuentas principalmente implica el establecimiento de controles y regulaciones que eviten un uso inapropiado de los recursos.

Desde esta última perspectiva tecnócrata, las evaluaciones orientadas a la rendición de cuentas y, por tanto, a comprobar el grado de eficacia y eficiencia en el que los recursos han sido gestionados por sus responsables, suelen ser evaluaciones de tipo sumativas (Rossi et al. 2004). Los hallazgos de este tipo de evaluaciones a menudo son puestos a disposición de los equipos responsables de la toma de decisiones para facilitar el control y la fiscalización de los recursos de cara a ejercer una valoración final sobre la pertinencia del programa o proyecto en cuestión. La independencia e imparcialidad del equipo evaluador suele ser un requisito fundamental para la validez de las evaluaciones orientadas a la rendición de cuentas.

La mejora del programa: En contraste con el punto anterior, mejorar el objeto evaluado se constituye a menudo como el propósito fundamental de la práctica evaluativa. *“La evaluación aspira a guiar al programa evaluado en su perfeccionamiento y desarrollo”* (Vedung 1997, pág.110). Perfeccionar el programa o proyecto evaluados antes de su finalización está tradicionalmente ligado en la literatura a las evaluaciones formativas (Rossi et al 2004, Chelimsky 1997, Patton 2008).

El destinatario de este tipo de evaluaciones es habitualmente el propio equipo de gestión del programa o proyecto evaluado. Generalmente, en evaluaciones formativas, el equipo evaluador trabaja de manera cercana a los gestores del objeto evaluado, tanto en el diseño de la evaluación como en el desarrollo de la misma y en la puesta en marcha de sus recomendaciones. En este sentido, la comunicación entre los equipos de evaluación y de gestión del programa es frecuente y tiene un carácter más informal que en la evaluación destinada a la rendición de cuentas (Rossi et al. 2004).

La generación de conocimiento: En ocasiones, la práctica evaluativa se distancia de la casuística particular de los objetos evaluados y persigue inferir sus hallazgos en disciplinas epistemológicas de interés para instituciones y organismos que las desempeñan. Estas evaluaciones son consideradas “*investigación fundamental que persigue incrementar el conocimiento general de una realidad determinada*” (Vedung 1997, pág. 110).

En un contexto organizacional, el gran desafío de las evaluaciones que persiguen este propósito, estriba en su capacidad de transformar sus hallazgos en conocimiento habilitando así a los destinatarios de la evaluación para tomar acciones adecuadas.

Nonaka y Takeuchi (1999) consideran que el conocimiento, aunque similar a la información en algunos aspectos, mantiene diferencias sustanciales con ésta última. Entre ellas, se destaca que mientras que la información en sí misma es a menudo factual y carece de propósito, el conocimiento va más allá y nutriéndose de información, ofrece una intención, una perspectiva y una postura sobre un área particular, la cual nos habilita a iniciar un curso de acción determinado para afrontar situaciones en áreas particulares.

Tal y como Patton (2002) sugiere, una de las grandes lecciones de la última década es que gestionar información no es lo mismo que gestionar conocimiento. La gestión del conocimiento es un proceso colaborativo que persigue lograr resultados tanto en productividad como en innovación. En contraste, la gestión de la información se limita a recopilar, procesar datos. La gestión del conocimiento incrementa nuestra capacidad para la toma de decisiones adecuadas (Executive Edge Journal 1998).

De acuerdo con su naturaleza racional, el conocimiento, según describe Palacios-Maldonado (2000), puede ser clasificado en diferentes categorías. Existe así: i) el conocimiento proposicional orientado a comprender el objeto de estudio; ii) el conocimiento procedural, enfocado en los factores y procesos que describen el medio para alcanzar un fin; iii) el conocimiento del descubrimiento, centrado en las raíces de los acontecimientos estudiados y finalmente; y iv) el conocimiento contextual que, como su nombre indica, contribuye a entender los aspectos que rodean al objeto de estudio. Cada una de estas categorías se transmite de un individuo a otro a través de diversos medios y es precisamente su transferibilidad la característica que hace del conocimiento un objeto de aprendizaje.

En relación a este último punto, una parte muy importante de la literatura especializada coincide en afirmar que el elemento clave para definir la transición entre conocimiento y aprendizaje es el cambio (Moxnes 1981). Existe, por consiguiente, un cierto consenso en relación a la idea de que el aprendizaje es un proceso a través del cual el sujeto modifica su comportamiento y desarrolla o adquiere nuevas formas de actuación (Palacios - Maldonado 2000). El nuevo conocimiento así generado se incorpora al ejercicio lógico de cada persona y modifica sus reglas de decisión y conducta. La capacidad de aprendizaje se desarrolla cuando las personas crean una nueva

mentalidad, cambian la manera de entender las cosas y afrontan las dificultades de una manera distinta.

El aprendizaje organizacional, o la capacidad de un organismo de cambiar su 'modus operandi' para afrontar nuevos retos, es más factible en aquellas organizaciones que persiguen generar nuevos conocimientos, con frecuencia a través de la práctica evaluativa (*Learning Organizations*). Tal y como resume David Garvin (1993): "*Una organización que aprende es una organización capaz de establecer, adquirir y transferir conocimiento y, por tanto, capacitada para modificar su comportamiento en base a nuevo conocimiento adquirido*" (Garvin 1993, pág. 51).

Las evaluaciones orientadas a la generación de conocimiento son frecuentemente llevadas a cabo por agencias gubernamentales o asociaciones afines, vinculadas a la investigación de problemas sociales con el objetivo de proponer nuevos enfoques o interpretaciones sobre fenómenos de interés. Este tipo de evaluaciones, puesto que están dirigidas a hacer contribuciones al conocimiento científico y social así como a generar una base importante para la innovación, suelen desarrollarse con métodos de investigación rigurosos, alineados con estándares académicos. Sus hallazgos se diseminan en revistas científicas acreditadas y, en general, en medios relacionados con el mundo académico (Rossi et al. 2004).

Motivos estratégicos: En ocasiones, las evaluaciones pueden perseguir fines utilitaristas que responden a motivos estratégicos de la gerencia de una organización. Cuando este propósito predomina, nos encontramos ante evaluaciones que esconden a menudo información relevante del programa evaluado para sobresaltar una imagen preconcebida del mismo (al servicio del motivo estratégico que persigue la evaluación) y en general dar una apariencia disfrazada de la realidad que permita justificar, en definitiva, una serie de acciones determinadas (Vedung 1997). En ocasiones, los motivos estratégicos, o también la existencia de 'agendas ocultas' como las denomina Rossi et al. (2004), persiguen que se utilicen los hallazgos de una evaluación para fundamentar y justificar una opinión arbitraria no necesariamente basada en evidencias y las acciones que de la misma se deriven.

Comúnmente, este tipo de uso, también llamado simbólico, se produce cuando bien los responsables políticos, los equipos directivos y/o el personal involucrado en el programa persiguen objetivos predefinidos sobre el futuro del objeto de estudio y utilizan a la evaluación para legitimar su decisión e incrementar adherentes a la misma. Es decir, la evaluación se convierte en un instrumento de legitimización y persuasión.

Cuando el propósito último de la práctica evaluativa responde a motivos estratégicos, el equipo evaluador se puede encontrar ante un problema ético al incrementarse sustancialmente el riesgo de que su integridad profesional se vea comprometida (Rossi et al. 2004)<sup>8</sup>.

---

<sup>8</sup>Los motivos estratégicos son considerados ejemplos de mala utilización. Más adelante se explica en detalle este concepto (apartado 1.5).

## 1.2. La utilización en evaluación

En el apartado anterior se relatan los diferentes propósitos de la práctica evaluativa, entre los que destacan la rendición de cuentas, la mejora del programa evaluado y la generación de conocimiento. Estas finalidades que, en ocasiones, se presentan de forma combinada, son todas ellas importantes y beneficiosas para los grupos de interés del programa evaluado en particular.

Sin embargo, además de definir con claridad el propósito de la práctica evaluativa, es fundamental analizar si los resultados que ésta arroja son realmente utilizados por los destinatarios de la evaluación. El uso de las evaluaciones ha generado diversas teorías y ha sido objeto de amplia investigación e intenso debate. De hecho, probablemente el uso de la evaluación sea el área donde más se ha investigado en la literatura sobre evaluación (Christie 2007; Henry y Mark 2003a).

La mayoría de las investigaciones realizadas sobre la práctica evaluativa destinan importantes esfuerzos a analizar aquellos factores que facilitan el uso y la aplicación de las recomendaciones que emanan de las evaluaciones (Preskill y Caracelli 1997). Incluso para algunas evaluadoras como Carol Weiss éste es un aspecto fundamental: "*La utilización lo es todo en evaluación*" (Weiss 2004, pág. 161). Tanto es así, que otros autores y autoras consideran que las evaluaciones deben ser juzgadas en base a la utilización que se haga de ellas (Patton 2008; Pawson y Tilley 1997; Torres et al. 1996).

Sin embargo, analizar la utilización de las evaluaciones no es un tema sencillo de abordar. Los matices han cobrado importancia en esta área de investigación, donde incluso existe divergencia de opiniones en cuanto a si los términos 'uso' y 'utilización' poseen el mismo significado. De esta forma, la terminología empleada en evaluación ha sido en sí misma un importante tema de investigación y reflexión, incluyendo los conceptos y términos que rodean al uso/utilización de las evaluaciones (Hopson 2000). A pesar de esta controversia, adelantamos ya que en este trabajo, se emplearán ambos términos de forma indistinta.

En este sentido, mientras Weiss (1980) abogó por emplear el término 'uso' al entender que éste mostraba de forma más explícita el carácter no episódico (no *ad hoc*) de las evaluaciones, Patton (2008), por su parte, considera que el término 'utilización' es más apropiado al reflejar mejor el valor agregado de los propios procesos inherentes de la práctica evaluativa. Saunders (2012) matiza, a su vez, que el término utilización incide en la intencionalidad teórica, la cual está implícita en el diseño de una evaluación, mientras que el término uso se enfoca en la capacidad institucional para poner en práctica los hallazgos, procesos o productos generados a través de la evaluación.

Más allá de la discusión anterior, la literatura recoge diferentes posibles usos o tipos de utilización en relación a los resultados de una evaluación. Por citar algunos, puede considerarse que las evaluaciones son utilizadas cuando se producen acciones que ocurren como resultado directo de ellas (uso instrumental), o cuando sus resultados han servido para aprender sobre el

programa, sus participantes, operaciones o efectos particulares (uso conceptual). A su vez, los usos de las evaluaciones pueden ser intencionales (normalmente se corresponden con los usos definidos en la fase de diseño de la evaluación) y también la evaluación puede generar usos no intencionales.

La responsabilidad en el uso de las evaluaciones ha sido también un elemento que ha ido evolucionando a lo largo de los años. Originalmente se consideraba que el uso de la evaluación era algo completamente desvinculado de las responsabilidades del evaluador/a, mientras que teorías más recientes se inclinan por un tipo de evaluador/a que tiene un rol más activo sobre el uso que se deriva de su trabajo. Además de este cambio en cuanto al rol del evaluador/a, las teorías más actuales sobre el uso de las evaluaciones consideran la necesidad de un replanteamiento del concepto del uso hacia otro concepto denominado 'influencia', donde se puedan incluir aspectos considerados hasta el momento de 'segunda categoría', como los efectos no intencionados de las evaluaciones y aquellos que evolucionan gradualmente a lo largo del tiempo (Kirkhart 2000).

Esta evolución hacia una mayor responsabilidad de la persona que realiza la evaluación ha derivado en la necesidad de entender y explicar los elementos que contribuyen a la utilización de la evaluación dentro del sistema de conocimiento de las organizaciones. Así, cada vez es más evidente que las evaluaciones en sí mismas como prácticas aisladas, tienen pocas posibilidades de generar cambios y de ser utilizadas, a no ser que la información generada se transmita adecuadamente. En definitiva, se trata de que las evaluaciones no sean concebidas como meros estudios aislados e independientes sino que formen parte de un sistema estructurado de aprendizaje prolongado en el tiempo.

Para comprender los términos de este debate conceptual y la evolución de la cuestión de la responsabilidad en el uso de las evaluaciones es preciso hacer una referencia específica, aunque breve, a la evolución de la práctica de la evaluación. Este repaso nos permitirá constatar como el uso de las evaluaciones se ha convertido en un tema de interés particular en la literatura especializada desde que la evaluación comenzara a ser considerada como una práctica de especial relevancia en la década de los 70. Así, a lo largo de los años, el significado del término uso o utilización de la evaluación ha evolucionado desde una perspectiva simple a una más sofisticada y multidimensional.

A partir de la segunda mitad de la década de los años 60, la práctica evaluativa adquiere especial relevancia en Estados Unidos, al constituirse ésta como una actividad a cargo del estado, orientada a valorar el rendimiento del presupuesto invertido en programas y servicios de índole social. Así, se puede afirmar que la demanda de la evaluación como actividad sistemática crece en paralelo al incremento de los programas gubernamentales de bienestar social (Shadish y Luellen 2005, House 1993). En particular, el gobierno estadounidense impulsó numerosos programas en diversos ámbitos como el de salud, vivienda y trabajo con el fin de reducir los niveles de pobreza existentes en aquel momento.

En la década de los años 70, el incremento de la inflación y de los impuestos producidos en el marco de la crisis del petróleo produjo un notable descenso en el desarrollo de los programas de carácter social, lo que disminuyó también la intensidad de la práctica evaluativa. Ante la escasez de recursos gubernamentales, las evaluaciones se convirtieron fundamentalmente en herramientas orientadas a justificar el uso de los recursos empleados con el propósito de rendir cuentas (*accountability*) y valorar el mérito de los programas evaluados. Así, las evaluaciones servían como instrumentos que guiaban la asignación de fondos en nuevos programas gubernamentales. En definitiva, se promovía en este período un uso instrumental de las evaluaciones.

En cuanto al rol del evaluador/a, existía una concepción generalizada de que la responsabilidad de los y las investigadores/as (todavía el término evaluadores no era muy empleado) se centraba únicamente en el diseño de los estudios de evaluación, la correspondiente recolección de datos y la publicación posterior de los hallazgos.

Así, se eximía a la persona profesional encargada de la evaluación de cualquier responsabilidad relacionada con el uso posterior de los resultados, responsabilidad que, por otro lado, recaía en el nivel de la gerencia o toma de decisiones, es decir, en las instituciones gubernamentales que aprobaban los programas de carácter social que eran evaluados. Los criterios para valorar la calidad de las evaluaciones en sí mismas se asentaban en las prácticas académicas vinculadas a las ciencias sociales, destacando la importancia que se otorgaba al rigor metodológico que permitía inferir conclusiones robustas y contrastables. En contraste, el grado de atención que se otorgaba al seguimiento realizado a los resultados de la evaluación era menor y fundamentalmente estaba excluido de los procesos anteriores.

En consecuencia, a finales de los años 70 emergen posiciones críticas ante la escasa utilización de los resultados y hallazgos que generaba la práctica evaluativa, poniendo así en evidencia la insuficiente capacidad de los métodos de investigación en sí mismos para ejercer cierta influencia en los procesos de toma de decisiones. Por el contrario, se hizo cada vez más patente el hecho de que las evaluaciones, para ser efectivas, requerían de algo más que simplemente generar información de calidad a través de prácticas de investigación adecuadas.

En este contexto, crece el interés por identificar los factores que inciden en un mayor uso de las evaluaciones así como por dilucidar las responsabilidades que cada parte interviniente tiene en las diferentes etapas del proceso evaluativo. Así, por primera vez, empezó a emerger cierto grado de conformidad con la idea de que los equipos de evaluación también podían 'rendir cuentas' por el uso de los resultados que las evaluaciones realizadas producían, de la misma manera que el personal a cargo de los programas evaluados debían rendir cuentas por la ejecución de sus propios programas. Es decir, se plantea una concepción de la evaluación que trasgrede los límites hasta entonces vigentes, convirtiéndose ésta en un proceso informativo continuo donde tanto los y las evaluadores/as como el personal del programa

asumen un grado de responsabilidad compartido en generar, transmitir y gestionar la información proveniente de la evaluación.

Este nuevo rol del evaluador/a que se extiende más allá de garantizar el rigor metodológico de la evaluación y que le otorga un cierto grado de responsabilidad por el uso de la misma, supuso un claro incremento del protagonismo de los y las evaluadores/as y, a su vez, la necesidad de establecer mecanismos que fiscalizaran su actuación. Así surgen los siguientes interrogantes: ¿quién evaluará a los y las evaluadores/as?, ¿bajo qué criterios se evaluará el grado de atención otorgado a la evaluación? Y, en caso de ser éste insuficiente, ¿cuál será el grado de responsabilidad del evaluador/a?

De forma simultánea, se produce un intenso debate entre las corrientes de evaluación denominadas 'experimentalistas' y 'naturalistas'. La primera se concentraba en juzgar los méritos de la práctica evaluativa a partir de criterios vinculados al rigor metodológico empleado, mientras que la segunda iban más allá y consideraba otros aspectos como el grado de participación de las diferentes partes intervinientes en el proceso evaluativo e, incluso, criterios de carácter ético, como el nivel de objetividad e imparcialidad del trabajo de evaluación y de sus hallazgos.

Este debate, así como el incremento del protagonismo del evaluador/a mencionado anteriormente, derivó en la necesidad de desarrollar unos estándares de evaluación que no sólo resolvieran las tensiones entre experimentalistas y naturalistas (Patton 2008), sino que, a su vez, sirvieran de referencia para delinear las responsabilidades y los roles de la profesión de la evaluación y del resto de las partes vinculadas con la práctica evaluativa.

Los años 80 representan, así, un período crucial para la conformación de la evaluación como disciplina profesional propia. En estos años, surgen agrupaciones de especial relevancia para la práctica evaluativa que contribuyeron al desarrollo de estándares de evaluación. Así, en 1984, la Sociedad de Investigación en Evaluación (*Evaluation Research Society*) y la Red de Evaluación (*Evaluation Network*), ambas de Estados Unidos, se unen para dar lugar al nacimiento de la Sociedad Americana de Evaluación (*American Evaluation Association*), la cual es hoy en día una de las instituciones referentes en materia de evaluación. Asimismo, en otros países también se establecen agrupaciones similares como la Asociación Canadiense de Evaluación y la Asociación Australiana de Evaluación.

En este período, y como ya se ha mencionado, uno de los temas que más controversia suscitaba era precisamente tratar de delimitar las responsabilidades de la profesión en relación con los posibles efectos y con el uso de las evaluaciones. Uno de los debates más importantes fue el mantenido por Weiss y Patton (Patton 1988a, Weiss 1988b, Weiss 1988c). Weiss argumentaba que los y las evaluadores/as, al operar en un contexto caracterizado por una amplia complejidad, podían únicamente aspirar a la generación de hallazgos sólidos y precisos que sirvieran de plataforma para dar respuesta a las preguntas planteadas por la evaluación en cuestión. Patton, por su parte, defendía una posición donde las personas profesionales responsables

de la evaluación detentaban un mayor grado de responsabilidad, al recaer sobre ellas una serie de obligaciones específicas vinculadas con la necesidad de promover y cultivar el uso de la investigación evaluativa. Para ello, Patton abogaba por que los y las profesionales de la evaluación asumieran responsabilidades orientadas a ayudar directamente a la toma de decisión, identificando con claridad y precisión las necesidades de información sobre el objeto evaluado, para poder, de esta manera, generar hallazgos que dieran respuestas adecuadas y completas, facilitando así el uso de la práctica evaluativa.

A pesar de las discrepancias existentes en cuanto a quién es responsable de facilitar y promover el uso de los hallazgos de la evaluación, a finales de la década de los 80, dentro de la comunidad evaluadora se alcanza un consenso importante sobre otra cuestión igualmente relevante: la necesidad y la importancia de analizar el contexto político y organizativo en el cual se desarrolla la práctica evaluativa, como un factor relevante para posibilitar el uso futuro de las evaluaciones. Diversos autores y autoras contribuyeron con sus estudios a poner de manifiesto la importancia del contexto en la práctica evaluativa y, en particular, en el uso y utilización de las evaluaciones.

En esta línea, Greene (1990), King (1995) y Weiss (1988c) defendían que las personas responsables de la toma de decisiones en un organismo determinado no son actores que actúan en solitario, ni toman decisiones de manera aislada e independiente sino que, por el contrario, sus decisiones son el resultado de un proceso deliberativo desarrollado en función de los grupos de interés a quienes representan. Además, la complejidad de los niveles burocráticos que existen dentro de las organizaciones, las líneas de comunicación entre dichos niveles y los mecanismos para canalizar la información surgida a partir de la práctica evaluativa se revelaban, también, como factores determinantes a la hora de utilizar los hallazgos de la evaluación (Preskill y Torres 1996).

En años posteriores, ya en plena década de los años 90, además de extenderse con fuerza esta idea de la evaluación como una actividad adicional dentro de un contexto organizacional, en vez de como una acción aislada, las evaluaciones empiezan a ser entendidas no solo como informes puntuales donde se detallan hallazgos relativos a los méritos del objeto evaluado, sino también como prácticas evaluativas que necesitan de un proceso duradero en el tiempo. En contraste con la década anterior en la cual el uso de la práctica evaluativa se había investigado de forma aislada e individual, en los 90 nace un interés particular en investigar el uso de la evaluación en las organizaciones como una práctica integrada dentro de las mismas.

De esta manera, la evaluación como herramienta para el aprendizaje organizacional emerge como una nueva utilidad de la práctica evaluativa, en paralelo con la importancia suscitada alrededor de la necesidad de investigar la utilidad y el uso de las evaluaciones. En este sentido, se elaboraron en esta época diversos estudios dirigidos a analizar los efectos de la evaluación en contextos organizacionales concretos (Jenlik 1994; Owen y Lambert 1995; Preskill y Torres 1996).

La fuerte inclinación durante estos años 90 a entender la evaluación como una actividad integrada dentro de un contexto organizativo dio pie a analizar la práctica evaluativa desde el prisma de la participación - en el proceso evaluativo - de las diferentes partes (personas, grupos) que interaccionan dentro de un organismo, emergiendo así los denominados enfoques participativos de evaluación. En contraposición a los modelos estrictamente consultivos como forma de elaborar evaluaciones, el enfoque participativo aspiraba a lograr una relación de equidad entre el equipo responsable de la evaluación y los actores clave dentro del organismo en el que se ubica el objeto evaluado.

En este sentido, se asienta con fuerza la idea de que cuando los equipos de evaluación trabajan conjuntamente con los actores clave dentro de las organizaciones, se incrementa sustancialmente la capacidad de los primeros de estimular la generación de conocimiento así como de favorecer la predisposición de los grupos responsables de la toma de decisión a considerar los hallazgos de la evaluación y a impulsar su utilización (Ayers 1987; Greene 1988a).

Así, durante los años 90, la participación de actores clave en el marco de las evaluaciones fue ampliamente promovida como un instrumento al servicio del fortalecimiento de las capacidades del organismo en cuestión y, en definitiva, del aprendizaje organizacional (Cousins y Earl 1992) y del empoderamiento de actores clave (Fetterman 1994). El auge de los enfoques participativos implicó así el surgimiento de una concepción del evaluador/a como un/a profesional que ofrece asistencia y apoyo dentro de las organizaciones. Shadish y Epstein (1987) explícitamente denominaron esta función/rol como 'orientación de servicio a los actores clave'. Este nuevo perfil del evaluador/a -orientado al cliente- ha sido ampliamente destacado por diferentes expertos/as de la comunidad evaluadora como una pieza fundamental que permite promover con éxito el uso de la evaluación (Cousins y Earl 1995; Owen y Lambert 1995; Preskill 1994; Torres et al. 1996).

Sin embargo, el modelo de evaluación participativa orientado al servicio presentó nuevos retos a la comunidad evaluadora. Este enfoque supuso un desafío para los/as evaluadores/as, puesto que el hecho de fomentar una estrecha relación con actores clave vinculados al objeto evaluado podía poner en peligro la objetividad de su juicio y capacidad evaluativa, así como facilitar la exposición y vulnerabilidad de su tarea profesional a posibles presiones institucionales. En este sentido, en la literatura especializada de los años 90, se sugerían algunas medidas preventivas para hacer frente a estos nuevos riesgos relacionados con la potencial pérdida de integridad del trabajo de evaluación. Entre ellas se destacaron prerrogativas importantes como, por ejemplo, la necesidad de disponer por parte del equipo evaluador de un conocimiento sustantivo del programa evaluado (Goering y Wasylenki 1993) y de su contexto político y organizacional (Cousins y Earl 1992, Greene 1988b) de cara a disminuir su vulnerabilidad a presiones externas.

La forma de mantener la integridad del trabajo de evaluación emerge, de esta forma, como una constante e importante preocupación (Scriven 1996) y ha

motivado que, en algunos casos, se haya propuesto limitar la aplicación de los enfoques participativos exclusivamente a evaluaciones formativas orientadas fundamentalmente a mejorar el desempeño de un programa o proyecto determinado durante el período de su implementación (Ayers 1987; Cousins y Earl 1992 y 1995). En cualquier caso, esta posible ruta para garantizar la neutralidad del trabajo de evaluación tampoco ha estado exenta de discusión. Fetterman (Fetterman 1994; Fetterman et al. 1996), por ejemplo, abogaba por no limitar la aplicabilidad de las evaluaciones participativas a las evaluaciones formativas. A través de las denominadas evaluaciones empoderativas, Fetterman propone que la responsabilidad relativa a la parte de la investigación que posea un carácter más técnico y especializado recaiga sobre los actores clave vinculados al proceso evaluativo, mientras que los/as evaluadores/as, por su parte, actuarían como guías o agentes facilitadores en todo el proceso, independientemente de que las evaluaciones en cuestión de carácter formativo o sumativo.

Los enfoques de evaluación participativa han generado, a su vez, una amplia corriente literaria centrada en desvelar las habilidades y capacidades idóneas de un 'buen/a' evaluador/a. Dentro de estas aportaciones, Ayers (1987) sugería que además de poseer una cierta destreza en el ámbito de la comunicación, los/as evaluadores/as requerían de habilidades sociales que facilitaran la creación de conexiones constructivas y distendidas especialmente con el personal vinculado al objeto evaluado que pueda tener cierta inclinación escéptica en relación a la potencial utilidad de la práctica evaluativa. Otras de las habilidades identificadas como significativas han sido las relacionadas con su función o rol como educador. La habilidad para transmitir conocimientos y para capacitar a personas adultas fue identificada como muy beneficiosa de cara a involucrar a actores clave en actividades de evaluación (Cousins 1996; Cousins y Earl 1992).

Como se puede apreciar en los párrafos anteriores, la utilidad de la práctica evaluativa ha representado un tema de extenso interés y debate durante varias décadas, interés que, por otro lado, ha estado ligado a la propia evolución de la evaluación como disciplina en sí misma y que ha dado lugar a un amplio desarrollo teórico dentro de la misma<sup>9</sup>.

### **1.3. Los tipos de uso de las evaluaciones**

De acuerdo a los modelos teóricos existentes, en este apartado se describen los tipos o categorías de usos tradicionales de las evaluaciones en función de sus resultados (Rich 1977; King y Pechman 1984; Leviton y Hughes 1981; Weiss 1988a). Así, los cuatro principales tipos de uso o utilización recogidos en la literatura evaluadora desde 1970 hasta la actualidad son: uso instrumental, uso conceptual, uso estratégico y uso de proceso<sup>10</sup>.

---

<sup>9</sup>La importancia del estudio de la utilidad queda reflejada en la contribución de Alkin (2004), que clasifica a la comunidad evaluadora (teoría y práctica) en función del énfasis que han puesto en: 1) los métodos de la evaluación; 2) su capacidad de asignar valor y 3) al uso de la evaluación.

<sup>10</sup>A efectos de armonización, en este trabajo se describen los cuatro tipos de uso comúnmente empleados en la literatura, aunque autores como Vedung (1997) proponen un mayor tipo de usos (7 en concreto): instrumental, conceptual, de legitimación, interactivo, táctico, ritual y como proceso.

Uso instrumental: Este tipo de uso fue el que originalmente se aplicaba cuando se comenzó a dar importancia al uso en las evaluaciones y hace referencia a acciones derivadas del resultado de las evaluaciones (de sus hallazgos) en el ámbito de la toma de decisiones, las cuales se convierten en cambios sustanciales en el programa evaluado. Es decir, la evaluación representa un instrumento de acción. La evidencia en este tipo de utilización se observa a través de decisiones y acciones directamente atribuibles a la evaluación como puede ser la aplicación de las recomendaciones o la aprobación de políticas gubernamentales vinculadas a las mismas. Así, fruto de la evaluación, los responsables de la toma de decisión pueden decidir dar fin a un programa determinado, permitir su continuidad, extenderlo en el tiempo o modificar partes sustanciales del mismo. Se espera, por tanto, que los responsables del programa evaluado usen las conclusiones y los hallazgos de la evaluación para tomar decisiones. El informe de evaluación es el producto principal de referencia de las evaluaciones de tipo instrumental.

Más en detalle, Weiss (1998) considera que el uso instrumental de las evaluaciones es factible bajo determinadas circunstancias: cuando las implicaciones de los hallazgos no son controvertidas ni van a provocar grandes conflictos en la organización; cuando los cambios son de pequeña escala y fácilmente 'digeribles' para la organización; y cuando el entorno del programa es relativamente estable y las recomendaciones emanadas del proceso de evaluación no implican grandes cambios en el presupuesto, en el liderazgo del organismo en cuestión o en la definición del tipo de clientes meta. Weiss también alude a una cuarta circunstancia que facilita el uso instrumental, la cual hace referencia a la particular situación en la que el objeto evaluado se encuentre sumido en una crisis importante o incluso paralizado y es difícil discernir el camino idóneo a seguir. En este caso extremo, la evaluación instrumental puede proporcionar una salida viable a la situación crítica particular en la que se encuentra el proyecto o programa evaluado.

Uso conceptual: Bien sea éste de aplicación interna o externa, el uso conceptual ocurre cuando una evaluación es capaz de influenciar de forma sustancial la forma en la que los actores clave vinculados al objeto evaluado piensan sobre el mismo (bien sean agentes internos de la organización o agentes externos) sin que necesariamente se produzcan cambios inmediatos o se asuman decisiones directas (Rossi y Freeman 1985). Con el paso del tiempo, el uso conceptual puede derivar en un uso instrumental y, con ello, originar modificaciones directas en el programa o proyecto en cuestión.

Cuando las modificaciones en este ámbito conceptual conducen a la generación de nuevo conocimiento en áreas que vayan incluso más allá del programa evaluado y que afectan a la esfera política en la que se enmarca el programa, nos encontramos ante una aplicación global de la evaluación, también denominada 'iluminación'. Así, de acuerdo con Weiss, 'iluminación' en este campo hace referencia a *“la percolación de nueva información, ideas y perspectivas que se producen en el marco de la toma de decisiones”* (Weiss, 1999, pág. 471). El conocimiento adquirido a través de la sistematización de la práctica evaluativa e incluso a través de las evaluaciones de impacto (impacto

entendido como efectos de largo plazo) puede generar este tipo de uso de tipo iluminativo (Rodríguez-Ariza 2014).

Uso estratégico: Hace referencia al uso que se produce cuando se utilizan los hallazgos de una evaluación de manera selectiva para justificar una posición u opinión - que previamente no estaba fundamentada en evidencias - con respecto a la necesidad de introducir cambios en el programa. Así en ocasiones, los equipos de gestión y el personal involucrado en el programa tienen claro qué cambios quieren hacer en el programa y utilizan la evaluación para legitimar su decisión y ganar adeptos. Es decir, la evaluación se convierte en un instrumento de persuasión<sup>11</sup>.

Otros tipos de utilización derivados del uso estratégico son:

- ✓ Uso simbólico: es el uso que se hace cuando la evaluación se realiza simplemente para aparentar un cierto grado de interés en datos e información sustentada relativa al objeto evaluado, aunque, en realidad no exista ninguna intención por implementar cambios en el programa o proyecto en cuestión.
  
- ✓ Uso de legitimación: es cuando los hallazgos de una evaluación son empleados para otorgar un cierto grado de crédito a una decisión que fue tomada antes de que se realizara la propia evaluación.

Es importante destacar que diversos trabajos (Greene 1988a y 1988b; Owen 1993) indican que los tres tipos de uso comúnmente aceptados por la literatura de evaluación que se pusieron fundamentalmente de manifiesto durante la década de los años 80 (instrumental, conceptual y estratégico) se encuentran de alguna manera interrelacionados y no son excluyentes entre sí. Es frecuente que los tres tipos de uso coexistan en una misma evaluación siendo el uso conceptual el que preceda a los otros dos tipos<sup>12</sup>.

Así, el uso conceptual en sí mismo tiene gran relevancia, en particular, para algunos autores (Leviton y Hughes 1981, Rossi et al 2004) que coinciden en la idea de que aun siendo el menos visible, el uso conceptual puede ser el más importante o 'poderoso', en el sentido de que su aplicación puede llegar a ser mucho más amplia que en los otros dos casos. Así, el uso conceptual en sí mismo, puede abarcar muchas formas en las que las evaluaciones pueden tener un efecto indirecto sobre las políticas y los programas evaluados. Estos efectos pueden extenderse desde la sensibilización a personas y a grupos sobre problemas y cuestiones actuales que atañen al objeto evaluado, hasta mejorar el conocimiento de métodos y prácticas exitosas que puedan ser aplicables en otros contextos y en otros programas.

---

<sup>11</sup>Este tipo de uso es considerado por muchos autores y autoras como una mala utilización de la evaluación. Más adelante se explica el concepto de mala utilización y sus tipos (apartado 1.4).

<sup>12</sup>Se entiende que para generar cambios operacionales (uso instrumental), antes se han tenido que producir cambios en la manera de pensar de las personas responsables de la toma de decisión, a través del uso conceptual.

Rossi cita, como ejemplo de lo anterior, los intensos debates públicos generados en Estados Unidos sobre el control de gastos en los servicios de salud, estimulados en buena medida por las evaluaciones efectuadas sobre la eficacia y eficiencia de los servicios sanitarios (Rossi et al. 2004).

Posteriormente, en la década de los 90, emergió con fuerza un cuarto tipo de uso denominado el uso de proceso. De hecho, para algunos/as autores/as el uso de proceso representa el tipo de uso de mayor relevancia en la actualidad (Cousins y Shulha 2006). Éste hace referencia a todo aquel cambio que se haya produce a nivel cognitivo, de comportamiento o incluso a nivel organizacional - generado de manera bien directa o indirecta - por encontrarse grupos de interés vinculados al objeto evaluado involucrados o inmiscuidos en el propio proceso de evaluación. Esto ha permitido que grupos clave dentro del organismo en cuestión hayan adquirido ciertas habilidades relacionadas con 'pensar de manera evaluativa' (Patton 2008). Así, el haber participado en una evaluación puede cambiar la forma de pensar y el comportamiento de las personas y, ello, puede a su vez derivar en cambios culturales y organizacionales (Preskill, Zuckerman y Matthews 2003). El uso de proceso ocurre cuando aquellos que han participado en la evaluación afirman que los cambios en el programa no han provenido de los hallazgos, sino de la transformación en el pensamiento de las personas relacionadas con el objeto evaluado, la cual ha estado originada durante el proceso de realización de la evaluación (Preskill y Caracelli 1996, Amo y Cousins 2007). Por su importancia para este análisis, este tipo de uso se desarrolla más en profundidad en los párrafos siguientes.

Como apuntan Preskill, Zuckerman y Matthews (2003), el uso de proceso encuentra su fundamentación en las teorías de aprendizaje constructivistas. Dichas teorías proponen que las personas construyen y desarrollan el conocimiento sobre una realidad compartida a través de su interacción y colaboración. En este sentido, la construcción, interpretación y validación de este nuevo conocimiento está circunscrita a un contexto determinado y parte de las experiencias que las personas han vivido sumidas en él (Bruner 1971; Campbell 1998).

El uso de proceso, que emerge en los años 90 en pleno desarrollo de la evaluación como disciplina profesional, aporta una nueva visión sobre el potencial beneficio que las evaluaciones pueden generar. Patton, en concreto, señala que tradicionalmente el uso que se había otorgado a las evaluaciones se había centrado en una concepción instrumental de las mismas y *"no se habían explorado maneras de conceptualizar o hablar de qué les ocurre a las personas o las organizaciones como resultado de estar involucrados en el proceso de evaluación"* (Patton 1998, pág. 225).

De hecho, los estándares del Comité Conjunto de Estándares para la Evaluación de la Educación, Comité mencionado anteriormente, definen evaluación como la investigación sistemática del valor o el mérito de un objeto (1981, 1994, 2011). Dicha definición, donde investigación y sistema se unen, implícitamente concibe la evaluación como un proceso, hecho del cual se deriva la importancia de la misma. En este sentido, Vedung (1997) sugiere que

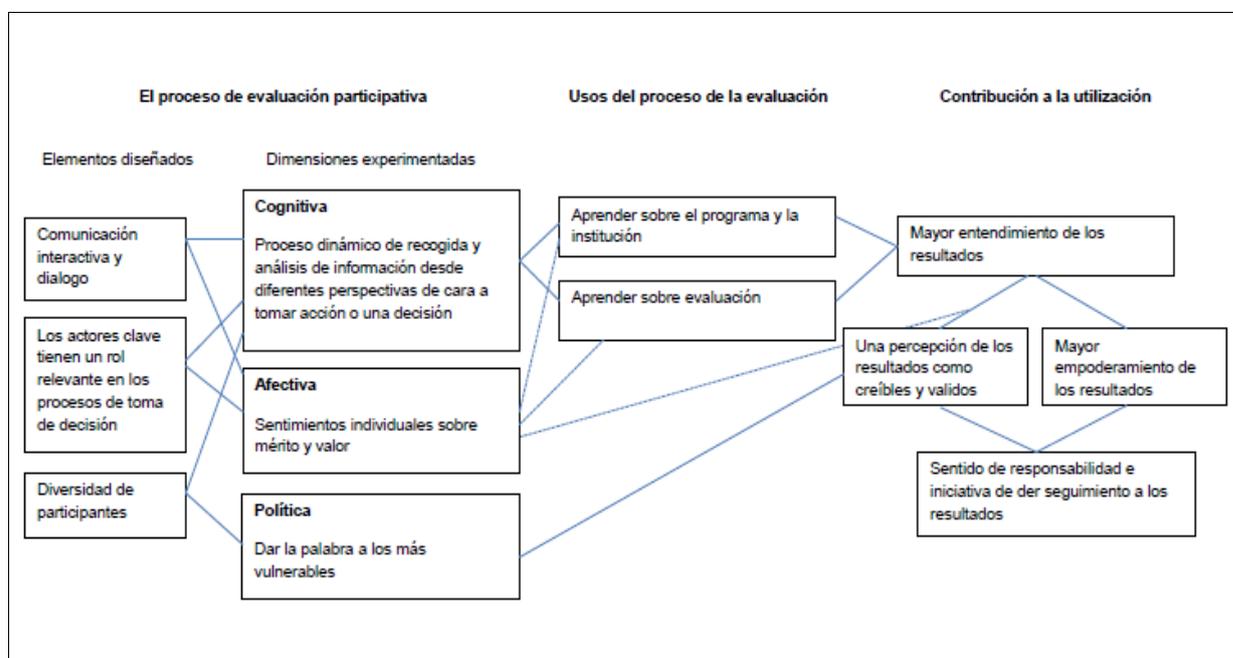
los enfoques tradicionales sobre el uso de la práctica evaluativa que confieren una gran importancia al informe final de evaluación como la única herramienta de utilidad, debe evolucionar a escenarios donde dicho informe no sea estrictamente el fin en sí mismo sino un producto más del proceso evaluativo.

El 'uso de proceso' trata así de diferenciar el aprendizaje que resulta del análisis de los hallazgos que son expuestos en el informe final de evaluación, del aprendizaje que se produce durante el proceso de evaluación, independientemente de lo que indique el informe final. Dicho aprendizaje tiene lugar debido a la participación de los actores clave vinculados al programa evaluado en dicho proceso de la evaluación.

Muchas de las contribuciones acerca del uso de proceso han relacionado el uso de proceso con el enfoque participativo de evaluación. Así, el uso de proceso es propio de las evaluaciones con carácter participativo, donde la colaboración e involucramiento de los actores clave facilita que permee en estos últimos el conocimiento evaluativo (Forss, Rebien y Carlsson 2002; Patton 1997 y 1998; Cummings 1999).

En 1988, Greene (1988a) contribuyó al desarrollo conceptual del uso del proceso a través de la diferenciación de tres dimensiones particulares que pueden verse afectadas por evaluaciones participativas que persiguen un uso de proceso: dimensión cognitiva, afectiva y política (ver Figura 1.1).

**Figura 1.1. Contribución del proceso de evaluación participativa a la utilización**



Fuente: Greene, 1988, pág. 108

La dimensión cognitiva hace referencia a aquellos cambios que se pueden producir en la forma de entender el programa, cambios que han estado estimulados por la discusión, reflexión y el análisis de las particularidades y de los problemas identificados durante el proceso de evaluación. Mientras que el

uso de proceso puede contener un componente instrumental (puesto que toda reflexión puede derivar en una decisión o acción específicas), este tipo de uso no se limita a cambios y medidas concretas, ya que la dimensión cognitiva de grupos clave vinculados al objeto evaluado puede verse afectada. Así, esta dimensión se enfoca en proporcionar un mejor entendimiento del programa por parte de los y las participantes por el mero hecho de estar involucrados en el proceso de la evaluación.

La dimensión afectiva está conectada a los sentimientos de las personas integrantes de los colectivos participantes durante el proceso de evaluación. Así, esta dimensión refleja los sentimientos individuales y colectivos que el proceso de evaluación genera como, por ejemplo, felicidad o tristeza por haber podido participar e influir de alguna manera en el proceso.

La dimensión política tiene lugar cuando el uso de proceso de la evaluación sirve para generar cambios sobre nuevas áreas de atención en el entorno de la toma de decisiones. Así, se considera que la evaluación participativa puede facilitar que colectivos considerados como más vulnerables puedan tener un papel representativo mayor en la toma de decisiones en las diferentes etapas de la evaluación.

#### **1.4. El mal uso en la práctica evaluativa**

Si en algo coincide la literatura sobre el uso de la evaluación es en afirmar la complejidad de desarrollar estándares sobre su incorrecta o mala utilización, fundamentalmente debido a la ambigüedad inherente al concepto del mal uso (Alkin 1990; Palumbo 1994).

En cualquier caso, a partir de la década los años 90, el mal uso o utilización de las evaluaciones se ha convertido en un importante tema de estudio en la teoría del uso de las evaluaciones. Así, el cambio de tendencia hacia los enfoques participativos ha generado serias preocupaciones acerca de la habilidad para mantener un grado suficiente de objetividad y no ceder ante las posibles presiones que puedan surgir del contacto con grupos de interés relacionados con el objeto evaluado, cuando se realiza un trabajo profesional de evaluación.

Existen diversos estudios que han conceptualizado las prácticas que implican una mala utilización de la evaluación (Alkin y Coyle 1988; Patton 1988b; Vroom, Colombo y Nahan 1994), y todos ellos coinciden en afirmar que generalmente existe un mal uso de la práctica evaluativa cuando ésta es manipulada, distorsionando sus hallazgos o corrompiendo los métodos de obtención de información. Así, la concepción más generalizada de mala utilización está vinculada con una reprochable intencionalidad previa, al implicar un acto calculado de suprimir, sub-representar o desbalancear los hallazgos de la evaluación para influenciar opiniones y/o decisiones preconcebidas (Patton 2008; Stevens y Dial 1994; Palumbo 1994). Sin embargo, también se distinguen otros tipos y ejemplos de mala utilización no intencionales:

✓ La mala utilización inadvertida, se produce cuando las personas que utilizan los hallazgos no poseen la competencia o experiencia adecuada para interpretarlos; no dedican tiempo suficiente a entenderlos; están más influenciadas por el estatus o la personalidad del evaluador/a que por sus hallazgos o simplemente no emplean la información de la manera más adecuada para su buen uso (Patton 2008).

✓ La sobre-utilización, se produce en aquellas situaciones en las que se presta demasiado énfasis a los hallazgos de la evaluación. Por ejemplo, resultados débiles son tratados como definitivos, o 'mejores prácticas' (*best practices*) son consideradas como rutas obligatorias a seguir (Patton 2008).

✓ El uso simbólico (ya mencionado anteriormente en el marco del uso estratégico), se produce en aquellas situaciones donde el acto de requerir una evaluación esta únicamente motivado por la voluntad de aparentar un cierto grado de interés por el objeto evaluado, a pesar de que no se tenga ningún interés ni por la evaluación ni por sus hallazgos (Duffy 1994; Stahler 1995).

Además de la mala utilización, un concepto paralelo es la 'no utilización' de una evaluación, la cual puede ser motivada por diversas circunstancias:

✓ La no utilización fruto de una mala evaluación (Patton 2005), también llamada la no utilización justificada (Cousins y Shulha 2006). Así, tanto Patton (2005) como Cousins y Shulha (2006) defienden que la no utilización 'consciente' puede ser una actitud apropiada, e incluso responsable, en aquellas situaciones donde exista una mala evaluación.

✓ La no utilización política, cuando los hallazgos son ignorados porque entran en conflicto con los valores, prejuicios, preferencias y predisposiciones del cliente/usuario y, por lo tanto, la evaluación es simplemente ignorada.

✓ La no utilización agresiva, en aquellos casos donde el cliente/usuario presenta un elevado grado de resistencia a los hallazgos generados y menosprecia activamente la evaluación y el trabajo realizado por los/as evaluadores/as (Patton 2008).

Siguiendo con las relaciones existentes, Patton (2008) también considera que se puede establecer un vínculo entre uso y mal uso. Así, contra más uso se hace de una evaluación, más posibilidades hay de que ésta sea mal utilizada, de la misma manera que el no uso implica también que no se realice mal uso. Sin embargo, esto no puede llevar a invitar a limitar el uso de la práctica evaluativa para no caer en malas utilizaciones o a proponer como objetivo

principal evitar exclusivamente el mal uso, ya que el hecho de que una evaluación no sea mal utilizada, no implica necesariamente que vaya a ser utilizada.

Finalmente, autores como Alkin (1990) proponen una serie de recomendaciones para prevenir la mala utilización en la práctica evaluativa, como, por ejemplo, buscar revisiones independientes del proceso evaluativo, revisar periódicamente aspectos metodológicos y consultar los códigos deontológicos existentes en el campo de la evaluación.

### **1.5. La investigación empírica existente sobre la utilización**

La evidencia empírica existente sobre los diversos factores que afectan al uso de las evaluaciones ha sido, en general, objeto de numerosas críticas basadas fundamentalmente en la escasa calidad de los datos aportados y, por tanto, en su limitada credibilidad (Leviton 2003; Owen 2006; Frey y Ledermann 2010).

Incluso entre los/as autores/as que han desarrollado marcos metodológicos detallados para profundizar en el uso de la práctica evaluativa así como en sus causas y efectos, se reconoce que las investigaciones empíricas existentes presentan importantes carencias: *“El estándar de evidencia que muchos de nosotros nunca aplicaríamos en la realización de evaluaciones, es el que frecuentemente predomina en el estudio del uso de las evaluaciones”* (Leviton 2003, pág. 526).

En la mayoría de los casos, las conclusiones provienen de la opinión expresada por diferentes personas y colectivos sobre la utilidad de la información de la evaluación, tratándose así de una valoración subjetiva sobre la cual no existe verificación alguna (Sanders 2002). La opinión acerca de si la evaluación ha sido de utilidad o sobre cuáles han sido los aspectos que han influido en dicha utilización no es susceptible de ser confirmada a partir de una estrategia de verificación sólida y constatable, tal y como sugieren diversas aportaciones a ese ámbito de estudio (Rich 1977; Leviton y Boruch 1983).

En muchos casos, la ausencia de verificaciones viene provocada por la imposibilidad de contrastar empíricamente los datos que confirman un uso determinado de la evaluación. Esto sucede generalmente cuando nos referimos a aspectos que pueden ser de naturaleza subjetiva como, por ejemplo, el desarrollo de análisis relativos al uso conceptual de la práctica evaluativa, donde se recogen únicamente percepciones personales ya que es ciertamente difícil encontrar una forma objetiva de comprobar hasta qué punto la práctica evaluativa ha supuesto una transformación en la forma de pensar de determinados grupos de interés relacionados con el objeto evaluado. En relación al uso conceptual de las evaluaciones, el más común de los tipos de utilización según Leviton (Leviton 2003), la verificación más allá de la opinión de las personas involucradas es ciertamente improbable y, por tanto, los estándares que se refieren a la evidencia necesaria para confirmar dicho uso han de ser hasta cierto punto poco exigentes. Esta idea es compartida por Henry y Mark (2003a, pág. 69): *“hay una falta de patrones de evidencia rigurosa y sistemática que pueda guiar la práctica de la evaluación o que los*

*evaluadores puedan usar para reflexionar o para mejorar su próxima evaluación”* (Henry y Mark 2003a, pág. 69).

En contraste, es más factible identificar medidas de verificación sobre la utilización de la evaluación cuando ésta es de tipo instrumental. El uso instrumental supone normalmente cambios verificables en el objeto evaluado y estas modificaciones pueden ser identificadas con relativa facilidad a través de diferentes métodos objetivos de recogida de información.

Para ahondar en esta cuestión con cierto criterio, a continuación se analizan dos ejercicios de investigación empírica que, a pesar de sus limitaciones, han servido de referencia en la literatura evaluadora para verificar los factores que afectan al uso de las evaluaciones. Estos dos estudios son la síntesis de Cousins y Leithwood (1986) y la adaptación posterior realizada de este primer estudio por Johnson et al. (2009).

Síntesis de Cousins y Leithwood en 1986: Este estudio es considerado por muchos como la revisión empírica más ambiciosa y rigurosa sobre el uso de la evaluación desarrollado hasta el momento (Johnson et al. 2009). Para la realización de esta investigación, sus autores seleccionaron una muestra de 65 estudios empíricos que habían sido llevados a cabo entre 1971 y 1985 sobre el uso de las evaluaciones, la mayoría de ellos basados en encuestas en las que se solicitaba a las personas encuestadas que enunciasen los factores que consideraban que habían facilitado el uso de los hallazgos de las evaluaciones ya realizadas. Para seleccionar la muestra de estudios empíricos utilizaron un filtro informático con palabras como ‘uso de las evaluaciones’, ‘uso de información’, ‘toma de decisión’, o ‘utilización del conocimiento’. Además complementaron este proceso con una búsqueda manual en revistas especializadas sobre el uso de las evaluaciones.

Los factores que determinaban el uso en estos 65 estudios empíricos fueron codificados en torno a 2 categorías:

- ✓ Categoría 1: La forma de realizar la evaluación (desarrollo de la evaluación)
- ✓ Categoría 2: El tipo de organización en la cual los hallazgos de la práctica evaluativa eran utilizados (contexto organizacional)

Cada una de estas 2 categorías contenía un total de 6 criterios. Así, bajo la categoría 1, los criterios eran: calidad de la evaluación, credibilidad, relevancia, calidad de la comunicación, hallazgos y oportunidad temporal de la evaluación. En la categoría 2 se incluían los siguientes criterios: necesidad de información, características de la decisión, clima político, información competitiva, características personales y compromiso o receptividad con/hacia la evaluación.

Estas categorías y sus criterios están recogidos en la Tabla 1.1.

**Tabla 1.1. Categorías y criterios del uso de la evaluación**

<b>Categoría 1: Desarrollo de la evaluación</b>	
Calidad de la evaluación	Características del proceso de evaluación, sofisticación de los métodos, rigor, tipo de modelo de evaluación utilizado.
Credibilidad	Objetividad, verosimilitud e integridad del proceso de evaluación y de las actividades del evaluador/a.
Relevancia	Grado de relevancia de la información proporcionada en la evaluación para las personas o equipos responsables de la toma de decisión.
Calidad comunicativa	Claridad y frecuencia del informe de resultados y grado de difusión/diseminación. También se incluye el nivel de aplicabilidad de las recomendaciones del informe y el proceso de comunicación entre evaluadores/as y clientes.
Hallazgos	Naturaleza de los hallazgos (positivos o negativos), nivel de congruencia con las expectativas de la audiencia, valor de los hallazgos para la toma de decisiones.
Oportunidad	Tiempo oportuno de la evaluación, puntualidad del informe una vez que la evaluación se ha finalizado, puntualidad en la diseminación de la información a las personas o equipos responsables de la toma de decisión.
Competencia del evaluador/a	Características personales en relación al nivel de adaptabilidad cultural y estilo de liderazgo.
<b>Categoría 2: Contexto organizacional</b>	
Necesidad de información	Incluye la necesidad de información de los usuarios de la evaluación, los tipos de información existentes, usuarios con necesidades de información distintas y necesidad percibida de realizar la evaluación.
Características de la decisión	El contexto sobre el cual se toman las decisiones, el tipo de decisión a realizar, la novedad del programa bajo estudio.
Clima político	Orientación política de las personas que solicitaron la evaluación. El grado de dependencia las personas o el equipo responsable de la toma de decisiones con respecto a fuentes externas. Restricciones presupuestarias.
Conflictos de información	Información relacionada con el objeto de la evaluación y accesibilidad a actores clave fuera del proceso de evaluación.
Características personales	Características del usuario de la evaluación, por ejemplo, el rol del usuario de la evaluación, el estilo de procesar la información, sus características sociales etc.
Receptividad a la evaluación	Actitud del usuario de la información hacia la evaluación y compromiso en la realización de la evaluación. El grado de resistencia de la organización a la evaluación. La amplitud de miras de los actores claves.

**Tabla 1.1. Categorías y criterios del uso de la evaluación (continuación)**

<b>Categoría 3: Participación de los actores clave</b>	
Receptividad	Grado de receptividad. Por ejemplo, hacer partícipes a diversos actores clave crea un mayor compromiso o receptividad hacia la evaluación.
Comunicación	La participación crea unos vínculos mayores de comunicación entre las personas responsables de la evaluación y las responsables del programa.
Relación directa	Facilita la comunicación y proximidad. Evita aspectos burocráticos.
Credibilidad	Facilita la credibilidad en el proceso de evaluación o del evaluador/a.
Hallazgos	Hacer partícipes a los actores clave para que conozcan los hallazgos de la evaluación y los comprendan y así se promueva el uso de la misma.
Aspectos organizacionales	Preguntar y buscar la perspectiva de los actores clave sobre aspectos organizacionales para que éstos se sientan partícipes de la evaluación.
Gestores	Buscar la participación de las personas responsables de la gestión del programa agiliza la toma de decisiones y facilita el uso de la evaluación.
Diferente tipos de tomadores de decisión	Tratar de involucrar a diferentes tipos de actores clave en función de los tipos de decisiones que sean necesarias para mejorar el uso.
Oportunidad	Involucrar a actores clave, facilitar que la evaluación responda a las necesidades de información de los actores clave y generar, por tanto, información clave para el uso.

Fuente: Elaboración propia, adaptada de Cousins y Leithwood (1986) y Johnson et al. (2009)

El estudio construyó un índice de prevalencia que permitía concluir, a partir de los estudios de la muestra, que la calidad de la evaluación era el criterio considerado por las personas encuestadas como el más importante para facilitar el uso de una evaluación. Esta conclusión es reflejo, en gran medida, de los criterios teóricos que imperaban en la década de los 70 a la hora de valorar las evaluaciones, criterios donde se destacaba la relevancia de la calidad técnica y del rigor metodológico de las mismas.

Ya se ha mencionado anteriormente que, en la década de los 80, se produjo una evolución muy importante del propio concepto del uso de la evaluación. En lineamiento con Shulha y Cousins, esta evolución se caracterizó principalmente por una mayor consideración del contexto como un factor crítico para entender el uso de las evaluaciones, además de destacar el uso de proceso como uno de los usos más significativos de la evaluación. También en este período se produce una expansión de la concepción de uso de las evaluaciones que transita desde una perspectiva individual hacia una visión organizacional, entendiendo el rol del evaluador/a como un agente facilitador, planificador y educador que desempeña sus funciones en el marco de la práctica evaluativa (Shulha y Cousins 1997).

En aras a reflejar los cambios producidos en el desarrollo del uso de la práctica evaluativa como concepto en sí mismo, Johnson et al. (2009) realizaron una nueva investigación empírica que utilizaba como base referencial los criterios

identificados en el estudio de Cousins y Leithwood de 1986, agregando la participación de los actores clave como una categoría adicional e incluyendo en consecuencia una mayor flexibilidad en el concepto de uso de las evaluaciones (ver Tabla 1.1).

Síntesis de estudios empíricos por Johnson et al. en 2009: Al igual que Cousins y Leithwood, para la selección de la muestra se realizaron inicialmente búsquedas electrónicas en revistas especializadas vinculadas al campo de la evaluación de los términos 'utilización de la evaluación' y 'uso de la evaluación', a los que se añadieron el término 'influencia de la evaluación'. Esta búsqueda electrónica fue complementada con búsquedas manuales en diversas bases documentales de publicaciones de evaluación.

Una vez preseleccionada la muestra, las publicaciones recopiladas debían además cumplir los siguientes requisitos para pasar a formar parte de la muestra final:

- ✓ Incorporar información sobre la metodología utilizada para la obtención de datos.
- ✓ Ser evaluaciones de programas o políticas.
- ✓ Estar publicadas.
- ✓ Incluir el uso de la evaluación o su influencia como una variable del estudio.
- ✓ La publicación debía de estar hecha entre enero de 1986 y diciembre de 2005.

Con la aplicación de estos criterios de selección, se pretendía excluir a todos aquellos estudios en los que no se podía verificar su rigor metodológico o que simplemente no seguían los estándares evaluativos. Como filtro final, la muestra obtenida fue refinada de acuerdo a un criterio de calidad adoptado por Guarino et al. (2004) y Guba y Lincoln (1989) que incluía aspectos como la rigurosidad de los métodos de investigación y la rigurosidad de las conclusiones basadas en evidencias, entre otros. Como resultado se seleccionaron 41 estudios para su posterior análisis.

Como se ha mencionado anteriormente, el nuevo marco adaptado de Johnson et al. (2009) incorporaba varias modificaciones respecto del anterior. El nuevo marco incluía una nueva y tercera categoría donde se hace referencia a la participación de los actores clave en el proceso evaluativo. Además, dentro de la categoría 1 (desarrollo de la evaluación), se incorporaba un nuevo criterio relativo a la competencia del evaluador/a. Estos cambios se han incluido en la Tabla 1.1, por lo tanto, la información recogida en esta tabla integra tanto las aportaciones realizadas por el trabajo original de Cousins y Leithwood en 1986 como las nuevas aportaciones realizadas por Johnson et al. (2009).

Las conclusiones del estudio empírico de Johnson et al. (2009) corroboran la teoría de que la participación de los actores clave favorece positivamente el uso de la evaluación, identificando dos elementos clave en este sentido: por una parte, en el propio proceso de la evaluación, es decir, durante el tiempo en que se lleva a cabo la evaluación, la participación de los actores clave mejora

el propio desarrollo de la evaluación (se cumplen los plazos, se mejora la calidad de los instrumentos de recogida de datos, etc.). Por otra parte, la participación ayuda a la posterior toma de decisiones en base a la información generada por la evaluación. El estudio sugiere que existe una relación entre participación e impacto de la evaluación (uso en el largo plazo). Como señalan Johnson et al. (2009, pág. 389), *“la implicación, interacción y la comunicación entre evaluadores y gestores del programa es importante para maximizar el uso de la evaluación en el largo plazo”*.

### **1.6. Debates conceptuales recientes: de la utilización a la influencia**

La década de los años 90 es considerada como una época de estancamiento teórico con respecto a la investigación sobre el uso potencial de las evaluaciones (Cummings 2002). Tanto es así que en algunos trabajos se mostraba un cierto pesimismo acerca del poco avance logrado en la evolución conceptual con respecto al uso y a la utilidad de las evaluaciones (Carlsson et al. 1999; Rebien 1996 y Thoenig 2000). A este pesimismo también se sumaron las dificultades encontradas en los escasos y limitados estudios empíricos existentes (como los mencionados anteriormente) que, a pesar de corroborar la relación entre ciertos factores y el uso de las evaluaciones, no habían podido establecer jerarquías entre los diferentes factores de forma rigurosa, destacando aquellos que poseen una mayor relevancia al presentar más capacidad para favorecer el uso (Ledermann 2012).

En este clima de cierta decepción y estancamiento teórico en el avance conceptual del uso de la práctica evaluativa, emerge a inicios de la primera década del nuevo milenio, un renovado interés por retomar la discusión teórica y tratar de superar las limitaciones del concepto de uso de las evaluaciones desde una perspectiva diferente (Caracelli 2000; Henry 2002; Kirkhart 2000; Ledermann 2012). El cambio de perspectiva estriba en abandonar la búsqueda de aquellos factores que facilitan el uso para centrarse en una nueva ruta teórica que busca profundizar en los mecanismos contextuales que desencadenan el uso (Henry y Mark 2003a y 2003b; Mark y Henry 2004).

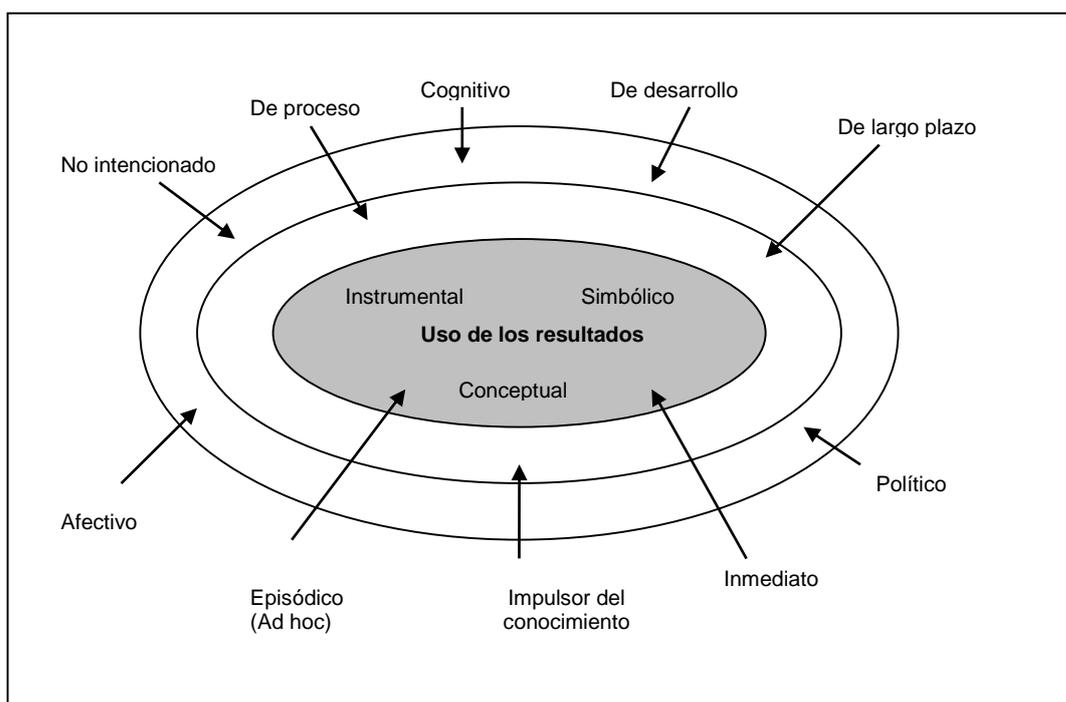
Así, Henry y Mark (2003a) identificaron como una de las principales limitaciones existentes en la evolución del concepto del uso de las evaluaciones, la falta de profundidad en el análisis de los tipos de efectos y cambios que una evaluación puede generar en su entorno. Por su parte, Kirkhart (2000) sugiere que la terminología relacionada con el concepto de uso utilizada en la década previa había sido demasiado rígida y estrecha, impidiendo captar una concepción más global y de mayor alcance en relación al impacto de las evaluaciones. En consecuencia, para reflejar y superar estas carencias, sugieren modificar el concepto de uso de las evaluaciones y sustituirlo por un nuevo concepto denominado de 'influencia' de la práctica evaluativa.

Kirkhart (2000) ha sido especialmente crítica con la noción tradicional de uso de las evaluaciones utilizada tradicionalmente, al considerarla una noción limitada y restrictiva debido fundamentalmente a su carácter unidimensional y circunstancial, el cual, en gran medida, dirige la atención relativa al uso de la

práctica evaluativa hacia los resultados de la misma. En su lugar, la autora propone un replanteamiento integral de la ruta a seguir para valorar de forma apropiada el impacto de una evaluación. Para ello, sugiere adoptar un nuevo marco teórico vinculado al concepto de influencia, entendida ésta como “la capacidad o el poder de las personas o las cosas de producir efectos en otros de manera intangible o indirecta” (Kirkhart 2000, pág. 7). Esta nueva teoría de influencia de Kirkhart ha encontrado múltiples apoyos (Christie 2007; Henry y Mark 2003a), sin embargo, también existen opiniones críticas que, a pesar de destacar su valor teórico, cuestionan su limitada aplicabilidad práctica debido a la dificultad de predecir y controlar aspectos relacionados con el impacto a largo plazo y los efectos no intencionados de las evaluaciones (Patton 2008; Alkin 2005).

En la Figura 1.2 se describen los posibles tipos de influencia identificados por Kirkhart, los cuales, desde su perspectiva de análisis, no han sido adecuadamente representados en la concepción tradicional de uso.

**Figura 1.2. Expansión del concepto de uso de la evaluación**

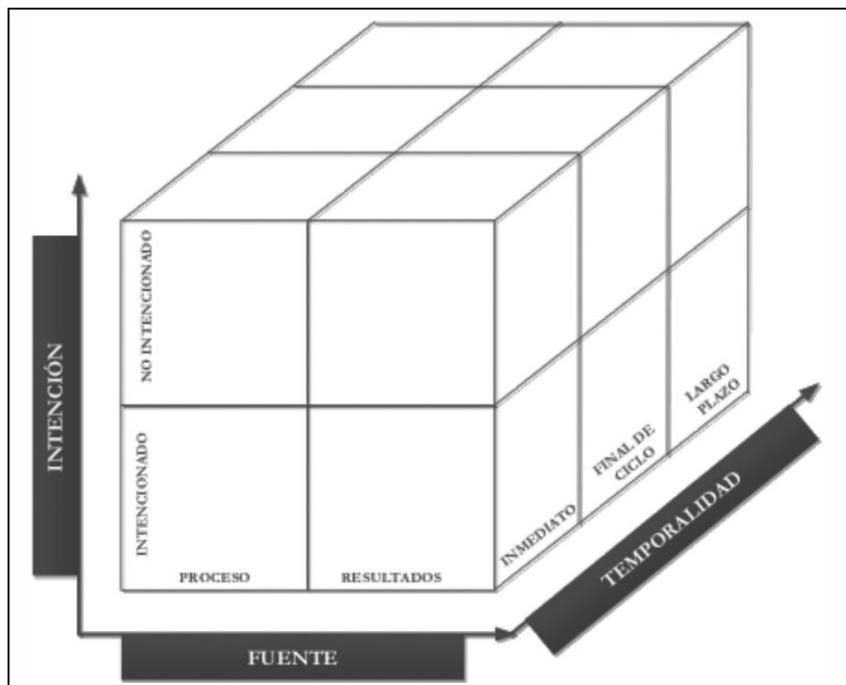


Fuente: Elaboración propia, adaptada de Kirkhart 2000, pág. 6

Para Kirkhart, el significado de influencia es más amplio que el de uso ya que al investigar la influencia se puede crear un marco conceptual que permita analizar efectos de las evaluaciones que son multidireccionales, incrementales, no-intencionales y no-instrumentales, además de los efectos tradicionales unidireccionales, episódicos, intencionales e instrumentales (y que están correctamente representados por el término uso).

La teoría integrada de influencia incorpora 3 dimensiones: la fuente de influencia, la intencionalidad y la temporalidad. Cada dimensión se subdivide, a su vez, en diferentes categorías, categorías que la propia autora reconoce como arbitrarias hasta cierto punto al representar variables continuas.

**Figura 1.3. Teoría integrada de influencia**



Fuente: Kirkhart, 2000, extraído de Monterde-Díaz 2014, pág. 28.

✓ La fuente de influencia: la primera de las dimensiones inherentes al concepto de influencia es definida como aquella que hace referencia a la fuente o al agente activo que provoca el cambio. Las dos categorías identificadas dentro de la fuente de influencia son el 'proceso' y los 'resultados' de la evaluación. A diferencia de los enfoques tradicionales de uso donde la separación entre proceso y resultado es manifiesta, Kirkhart plantea la posibilidad, bajo este nuevo enfoque, de que la fuente de influencia provenga de ambas categorías, es decir, del proceso evaluativo y de los resultados de la evaluación, y que ambas puedan estar interrelacionadas (Monterde-Díaz 2014).

✓ La intencionalidad: esta segunda dimensión se refiere al propósito que subyace en la práctica evaluativa y que, en definitiva, juega un rol importante en relación al tipo y grado de influencia que ésta pueda generar, bien sea a través del proceso de evaluación o de los hallazgos que emerjan. Esta dimensión, por primera vez, refleja la importancia de considerar que la influencia ejercida puede diferir de la intencionalidad original que motivó la práctica evaluativa, produciéndose así una influencia no intencionada que no debe pasar desapercibida por el evaluador/a.

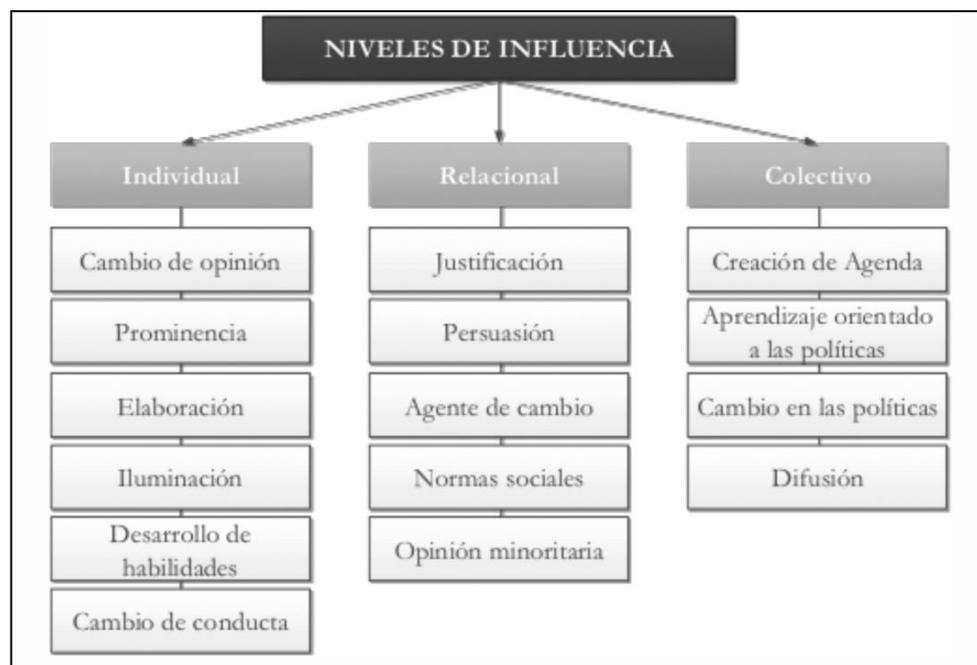
- ✓ La temporalidad: La tercera dimensión del concepto de influencia es la dimensión temporal en la que la influencia se produce y tiene lugar. Aquí, la autora diferencia 3 categorías: influencia inmediata, influencia al final del ciclo e influencia a largo plazo. De esta forma, se plantea un marco capaz de reconocer las diferencias entre la influencia ejercida durante el proceso de evaluación de aquella que ocurre posteriormente.

Según Kirkhart (2000), la teoría integrada de influencia presenta numerosas ventajas y contribuye a clarificar el debate sobre el uso de la práctica evaluativa al ilustrar mejor los diferentes efectos que la evaluación puede llegar a tener. A su vez, este nuevo marco teórico, a juicio de la propia autora, es capaz de registrar posibles tendencias a lo largo del tiempo, así como facilitar el meta-análisis de los estudios de evaluación a partir de las dimensiones anteriormente descritas.

Henry y Mark (2003b) coinciden con Kirkhart en las bondades de su nuevo enfoque vinculado a la influencia de las evaluaciones. En su opinión, el concepto de influencia ofrece una perspectiva analítica más completa y permite comunicar y proporcionar mejor las evidencias sobre las consecuencias que las evaluaciones pueden producir. Asimismo, este enfoque permite una mayor conexión entre la evaluación y su finalidad de mejora social al capturar aspectos que serían difíciles de entrever utilizando el enfoque tradicional. En consecuencia, los autores proponen un modelo de influencia en el que las consecuencias de las evaluaciones están necesariamente ligadas (positiva o negativamente) al propósito de mejora social que, según ellos, motiva en última instancia la práctica evaluativa.

Como parte de este marco, en su propuesta desarrollan una teoría de evaluación similar a la teoría de programas y clasifican los procesos de cambio y los resultados que una evaluación es capaz de generar en tres niveles: individual, interpersonal y colectivo. De esta forma Henry y Mark consideran a la evaluación como un programa más, con sus actividades, productos, resultados y objetivos, donde el objetivo de mejora social es considerado inherente a toda evaluación. A partir de este análisis, el marco teórico resultante permite realizar distinciones entre los cambios que las evaluaciones pueden generar en los tres niveles mencionados (individual, interpersonal y colectivo).

**Figura 1.4. Mecanismos a través de los cuales la evaluación produce influencias**



Fuente: Henry y Mark, 2003, extraído de Monterde-Díaz 2014, pág. 29.

El primer nivel, el **nivel individual**, se refiere a aquellos casos donde la evaluación (bien sea a través de los procesos inherentes a la misma o de los hallazgos que genera) provoca algún tipo de cambio, tanto en la forma de pensar como en el ejercicio u operación de una o varias personas, siempre a nivel individual. Los autores proponen seis mecanismos a través de los cuales se puede desencadenar este cambio a nivel individual:

- ✓ Cambio de opinión: la actitud u opinión de las personas responsables de la toma de decisión sobre un programa puede verse afectada debido a los hallazgos de una evaluación.
- ✓ Prominencia: la información originada por una evaluación sobre los efectos de un programa puede dirigir la atención a pensar en otros problemas sociales.
- ✓ Elaboración: un informe de evaluación puede estimular a prestar más atención al programa y a cuáles son las expectativas para producir buenos resultados.
- ✓ Iluminación: aspectos del programa que tal vez hayan sido obviados y que, después de la evaluación, pueden pasar a ser considerados importantes.

- ✓ Desarrollo de habilidades: nuevas habilidades adquiridas, como el aprendizaje de nuevas técnicas, fruto de la evaluación.
- ✓ Cambio de conducta: representa las acciones a nivel individual que se originan como resultado de la evaluación (por ejemplo, el personal docente de una institución educativa puede modificar su forma de practicar la docencia a partir de los resultados que arroje una evaluación sobre un programa educativo en particular).

El segundo nivel, el **nivel interpersonal**, se refiere al cambio generado en las personas de un organismo motivado por las interacciones que se han producido entre ellas durante el proceso de evaluación. En este nivel se identifican cinco diferentes formas de comportamiento interpersonal que pueden verse afectadas como consecuencia de una evaluación:

- ✓ Justificación. Se refiere al uso de una evaluación para justificar una posición ya adoptada previamente. Anteriormente, en la descripción de tipos de uso la citábamos como evaluación por motivos estratégicos.
- ✓ Persuasión. Una evaluación, generalmente a través de la comunicación directa, puede persuadir a una o varias personas o grupos para que modifiquen su comportamiento y/o actitud frente a determinadas prácticas o posiciones.
- ✓ Agente de cambio: se refiere a aquellas situaciones donde una evaluación puede llevar a alguien (persona o grupo) a erigirse como agente de cambio o líder de una determinada posición destinada a modificar políticas o procedimientos establecidos en el organismo en cuestión.
- ✓ Normas sociales: a través de la práctica evaluativa, es posible informar y difundir formas de acción y normas de comportamiento inherentes a diferentes grupos que en el futuro pueden ser adoptadas en el colectivo propio.
- ✓ Opinión minoritaria: la práctica evaluativa al tratarse de un ejercicio sistemático de obtención de información y análisis de la misma, a través de rigurosos métodos de procesamiento, posee un cierto grado de robustez y credibilidad, el cual puede servir para legitimar posiciones minoritarias o menos extendidas socialmente.

El tercer nivel, el **nivel colectivo**, se refiere a la influencia directa o indirecta que las evaluaciones pueden ejercer en las decisiones y en las prácticas de las organizaciones, ya sean éstas públicas o privadas. En este punto, la atención se enfoca sobre cuatro resultados que representan cambios en este nivel:

- ✓ Creación de agenda. Una evaluación puede poner el foco de atención sobre un tema específico e incluirlo en la agenda política de un organismo. Es decir, los hallazgos de una evaluación pueden ser un paso importante para generar un espacio de discusión que conlleve cambios organizacionales o políticos mediante la inclusión de temas específicos en la agenda de los órganos responsables de la toma de decisión de dicha organización.
  
- ✓ Aprendizaje orientado a las políticas. Las evaluaciones proporcionan información analizada y procesada que puede generar un conocimiento fundado sobre los elementos que necesitan ser revisados dentro de un organismo para la consecución de determinados objetivos organizacionales.
  
- ✓ Cambio en las políticas. Se refiere a los cambios en las políticas o directrices generales establecidas en una organización, motivados por la práctica evaluativa.
  
- ✓ Difusión. Se refiere a cambios experimentados en organizaciones como resultado de evaluaciones llevadas a cabo en otra entidad.

A modo de conclusión, la teoría e investigación empírica reciente muestra la evolución y las nuevas perspectivas del concepto de uso. La tendencia actual se inclina a considerar el uso dentro de una perspectiva compleja y multidimensional, donde el concepto de influencia emerge con fuerza. La temporalidad, la intencionalidad y la fuente son dimensiones que enriquecen la concepción tradicional del uso en este nuevo marco de influencia, aunque indudablemente pueden añadir todavía más complejidad a su verificación empírica.

Los estudios empíricos existentes sobre los factores desencadenantes del uso son escasos y recogen ciertas limitaciones metodológicas. La mayoría de los más representativos, como los mencionados hasta ahora (Cousins y Leithwood 1986, Johnson et al. 2009), son estudios donde el equipo investigador ha pre-seleccionado los factores vinculados al uso de las evaluaciones en base a ciertas consideraciones teóricas. Esto abre la posibilidad a que existan otros factores que aún no hayan sido suficientemente tomados en cuenta, y que hayan podido pasar desapercibidos hasta ahora.

### **1.7. La evaluación y su convivencia con otras fuentes de información**

Las dificultades de conceptualización teórica y demostración empírica del uso y utilización de las evaluaciones son dos elementos de gran importancia para la práctica evaluativa, razón por la que ambos han sido analizados y debatidos de manera profusa en la literatura especializada, tal y como se ha venido reflejando en este primer capítulo.

En cualquier caso, durante los últimos años, cada vez es más frecuente encontrar opiniones que sostienen la necesidad de abandonar esta idea tradicional dirigida a buscar e identificar cómo atribuir, completamente, a la evaluación los posibles cambios de conocimiento o de proceder de las personas y organizaciones, y empezar a recuperar 'el espíritu de humildad' que reclamaba Weiss (Birkeland et al. 2005) en alguno de sus artículos y que, sin duda, implica un entendimiento más modesto de los efectos y del impacto de la práctica evaluativa. Por consiguiente, cada vez va tomando más fuerza, la concepción de las evaluaciones como herramientas de aprendizaje que forman parte de un contexto más amplio, definido por el conocimiento organizacional, y donde la evaluación (su proceso y sus hallazgos) actúa junto a otros elementos que operan simultáneamente al desempeño de evaluaciones (Leeuw 2006).

Esta postura descansa sobre la idea de que las evaluaciones no son el único tipo de información que sirve en los procesos de toma de decisiones (Furubo 2006; Stadler Waerness y Ovrelid 2006). Hoy en día, los canales de información para la toma de decisiones son muy variados y las evaluaciones son un insumo más en el proceso decisional. Davies (1999, pág. 159) plantea que: *"Evaluación es hablar sobre cambios que afectan a personas y a organizaciones. Una evaluación no puede pretender brindar ese cambio por sí sola. Lo que sí puede es contribuir a la generación de conocimiento para mejorar el desempeño de los programas y el rendimiento organizacional"*.

En esta misma línea argumentativa, Perrin (2006, pág. 23) especifica: *"La planificación, la gestión y el desarrollo de políticas necesitan de canales de información, formales e informales, donde la evaluación representa solamente un tipo de fuente"*. Y Mathison (2005, pág. 304) a su vez clarifica que *"una evaluación aislada no puede proporcionar la información suficiente para entender las características de la organización ni tampoco como dichas características influyen los programas evaluados"*.

En este nuevo contexto, la evaluación no es considerada de forma individual sino como parte de un sistema de información amplio y donde el uso de la evaluación ha evolucionado a una concepción más multidisciplinar e intangible donde se reconoce explícitamente que la evaluación cohabita con otras muchas formas de información (Mayne 2007).

Esta nueva visión invita a indagar sobre cuáles son los elementos que contribuyen a generar conocimiento en el marco de las organizaciones así como sobre qué aportaciones puede realizar la práctica evaluativa en este campo. Ya se ha mencionado como el hecho de generar conocimiento como propósito de la disciplina evaluativa está relacionado con el uso conceptual de las evaluaciones analizado anteriormente. Este tipo de utilización de la práctica evaluativa, que se remonta a la década de los años 70, reconoce que las evaluaciones pueden influenciar de forma sustancial la forma de pensar de los miembros de una organización a través de la generación de conocimiento en áreas de relevancia.

En particular, la relación entre evaluación y conocimiento ha sido analizada en detalle por Shadish (1987), quien subraya que un mayor entendimiento sobre los programas y proyectos a través de la práctica evaluativa puede contribuir, de forma fundamental, a la generación de teorías de ámbito general capaces de producir cambios y transformaciones sociales. Scheirer (1987), en la misma línea, afirma que las evaluaciones que se mueven en la esfera del conocimiento pueden dar un mayor sentido no solo a la propia razón de ser de los programas evaluados, sino también a los medios más apropiados para alcanzar los fines propuestos. Estos esfuerzos están, en última instancia, orientados a la generación de conocimiento y aunque pueden ser muy beneficiosos para el objeto evaluado, van más allá de la efectividad de un programa en particular.

Más recientemente, la generación de conocimiento a partir de la práctica evaluativa ha abierto un nuevo debate teórico que persigue esclarecer las formas idóneas de analizar, procesar y canalizar la información para transformarla en conocimiento organizacional, así como los medios para conseguir que las evaluaciones se integren en dicho proceso. Para ello, se han identificado diferentes instrumentos como, por ejemplo, las meta-evaluaciones, las cuales pueden afectar el *modus pensandi* y el *modus operandi* de la organización, y llegar a formar parte de su 'stock' de conocimiento. De la misma manera, Weiss (1999) describe como el análisis de las lecciones aprendidas y de las denominadas buenas prácticas que provienen del campo de la investigación y de la evaluación, puede ejercer también cierto grado de influencia a niveles de abstracción mayores al de los propios programas evaluados, llegando incluso a contribuir de forma substancial al desarrollo de las políticas de la organización. Esta situación, donde la práctica evaluativa ejerce una elevada influencia dentro de la organización a partir de la generación de conocimiento especializado en áreas de interés, se asemeja al fenómeno de 'iluminación' anteriormente mencionado.

Sin embargo, la contribución que la práctica evaluativa puede realizar al conocimiento organizacional se extiende más allá de las aportaciones que parten del análisis y sistematización de un conjunto de evaluaciones. Así, cuando la práctica evaluativa forma parte integral del sistema de información que tiene la propia organización, su capacidad de contribuir a generar conocimiento organizacional se ve reforzada. La literatura consultada también hace referencia a las dificultades que afronta la práctica evaluativa de cara a trascender dentro del entramado organizacional cuando su desempeño se realiza de forma aislada.

En aras a analizar en detalle los beneficios que para un organismo determinado pueda tener el hecho de que la práctica evaluativa forme parte integral del sistema de información organizacional, en vez de desempeñarse como una función aislada, es necesario primero identificar y describir los demás elementos que componen el sistema de información mencionado, es decir, aquellos mecanismos de información que conviven con la evaluación. En la actualidad, el monitoreo de programas y proyectos, derivado de la teoría gerencial denominada 'gestión por resultados', es el principal foco de información dentro de un organismo, junto con la práctica evaluativa.

La **Gestión por Resultados** surgió en la década de los años 90 precisamente como consecuencia, entre otros factores, de algunas de las limitaciones observadas en la función evaluativa. Así, muchos estudios coinciden en afirmar que el auge teórico y conceptual de la Gestión por Resultados está fuertemente influenciado por las limitaciones observadas en las evaluaciones como instrumentos útiles para la toma de decisiones. La búsqueda de nuevos enfoques que permitieran hacer frente a dichas limitaciones ha impulsado la aplicación y la extensión de la gestión por resultados. Bastoe justifica, de la siguiente manera, la enérgica emergencia de la gestión por resultados, fruto de la búsqueda de nuevos enfoques gerenciales: “*existe una gran desilusión con la práctica de la evaluación ya que muchos gobiernos experimentan un uso limitado de las evaluaciones. Los hallazgos de una evaluación son difícilmente incorporados en los procesos de decisión*” (Bastoe 2006, pág. 97).

Así, desde una perspectiva pragmática de gestión, se consideraba que los informes de evaluación frecuentemente producían un tipo de conocimiento de difícil aplicación ya que no siempre era explícito y, en numerosas ocasiones, se producía de forma tardía. “*La información debe llegar en el momento adecuado y esto es un problema de la evaluación en la sociedad del conocimiento*” (Nielsen y Ejler 2008, pág. 172). La falta de puntualidad, o mejor dicho de sincronización de la evaluación con las decisiones de asignación de recursos, era un limitante importante de la evaluación en una época en que gobiernos y organizaciones se ven obligadas a reformar los procesos presupuestarios en base a resultados (Davies 1999; Young 2006).

En efecto, en la década de los años 90, muchos países integrantes de la OCDE llevaron a cabo reformas del sector público en un intento de dar respuesta a importantes cambios políticos, económicos y sociales. Por citar algunas de las transformaciones y tensiones que tuvieron lugar en dicha época y que obligaron a los países desarrollados a implementar reformas y medidas estructurales de adaptación, se destacan, por ejemplo, las presiones de tipo económico estrechamente relacionadas con los denominados desafíos competitivos derivados de la globalización. A su vez, las prescripciones económicas de índole neoliberal se encontraban inmersas en un contexto donde se distingue de forma progresiva una tendencia a la desconfianza en el sector público, tanto en el ámbito político como social, que lleva a exigir más transparencia y rendición de cuentas a los representantes gubernamentales.

Algunas frases propias de la época fueron ‘reinventando el gobierno’, ‘hacer más con menos’, y ‘demostrar valor por dinero’, las cuales son ilustrativas de los importantes cambios que se fueron sucediendo durante esta década de los años 90<sup>13</sup>, década en la cual, además de los países desarrollados, también los

---

<sup>13</sup> Frecuentemente, estas reformas fueron dirigidas a través de la legislación propia y de órdenes ejecutivas de los países miembros de la OCDE. Por citar algunos ejemplos, el *Government Performance and Results Act* de 1993 supuso la punta de lanza en la reforma del gobierno federal estadounidense. En el Reino Unido, la publicación del documento *Better Accounting for the Taxpayers' Money* en 1995, tuvo una respuesta inmediata del gobierno con respecto a la incorporación de reformas en el sistema presupuestario. La experiencia fue similar en Australia, con la introducción del sistema presupuestario *Accruals-based Outcome and Output Budgeting*.

países emergentes sometieron a importantes reformas a sus sectores públicos. (Rist 2006)<sup>14</sup>

A pesar de las variaciones existentes, se puede concluir que las reformas coinciden en varios aspectos, entre los que se incluyen:

- ✓ El enfoque en criterios de desempeño (efectividad, eficiencia, calidad de los servicios).
- ✓ Orientación al servicio y a la adecuación de los servicios a las necesidades.
- ✓ Mayor participación de la ciudadanía en el control de la gestión pública.
- ✓ Reforma de los procesos presupuestarios y de los criterios para la asignación de recursos.
- ✓ Aplicación en áreas públicas de prácticas de gestión procedentes del sector privado.

Las reformas enfatizan la necesidad de mejorar el desempeño y de poner en marcha una serie de mecanismos para medir la eficacia de los gobiernos en el logro de sus objetivos<sup>15</sup>. Estos esfuerzos se produjeron a través de la introducción de nuevos mecanismos que permitían, de forma clara y directa, medir la eficacia y eficiencia de la gestión en el logro de resultados dando lugar a los que hoy se denomina 'gestión por resultados'.

Davies (1999, pág. 151), conceptualiza la gestión por resultados como una serie de esfuerzos sistemáticos, iniciativas y procesos que poseen las siguientes características:

- ✓ Define desempeño en términos de resultados (productos, efectos, impactos)
- ✓ Establece un conjunto de niveles medibles de logros obtenidos (metas de desempeño, estándares)
- ✓ Determina el grado de consecución de los resultados propuestos en base a indicadores de desempeño que deben ser monitoreados de forma continua.

---

<sup>14</sup> Chile, Filipinas, Sri Lanka, India son algunos ejemplos.

<sup>15</sup> Search for Results, Public Management Practices, OECD 1997

- ✓ Aboga por un enfoque de rendición de cuentas para el logro de los resultados en función de los recursos utilizados.
- ✓ Basa las decisiones sobre la asignación de recursos en información veraz y objetiva sobre el desempeño.

Binnendijk resume este nuevo enfoque gerencial afirmando que la gestión del desempeño, también conocida como gerencia o gestión por resultados, “*es una estrategia gerencial que tiene como objetivo principal lograr cambios importantes en la manera en la que operan las instituciones a través del criterio central de mejorar el desempeño para lograr mejores resultados*” (Binnendijk 2000, pág. 2).

Se establece así a partir de las bases descritas una línea divisoria entre el enfoque gerencial de gestión por resultados (*Performance Management*) y lo que se denomina la función de seguimiento (*Performance Measurement*). Ambas nociones son complementarias aunque presentan una serie de particularidades que las diferencian y que no siempre resultan obvias.

La **función de seguimiento**, también conocida como función de monitoreo, está orientada a la provisión de información relevante y oportuna a la gerencia del programa o proyecto en cuestión y, por tanto, sus actividades incluyen el desarrollo de indicadores así como la recogida periódica de datos sobre los mismos, posibilitando, así, una medición exhaustiva de los avances y los resultados que progresivamente se van obteniendo durante la fase de implementación del programa en cuestión. La **gestión por resultados**, por su parte, presenta un enfoque más amplio y a un nivel más cercano a la toma de decisiones, ya que persigue que la información generada a través de la función de seguimiento o monitoreo se corresponda con la que es efectivamente demandada por la gerencia del programa y se use, en tiempo y forma, no solo para valorar los resultados obtenidos hasta el momento sino también para adoptar decisiones adecuadas que permitan mejoras en distintos ámbitos. Por tanto, la gestión por resultados incluye el establecimiento de procedimientos, mecanismos e incentivos que garanticen que el monitoreo se lleve a cabo de forma oportuna y que la gerencia lo incorpore en su proceso de decisión. En un sistema de gestión por resultados, el logro de resultados y la mejora continua van necesariamente ligados a la información y a la recogida de datos (monitoreo) y, por tanto, la función de seguimiento se considera parte integrante del proceso de gestión por resultados<sup>16</sup>.

Como se observa, la gestión por resultados (y como parte de ella el monitoreo) y la evaluación son dos ‘movimientos’<sup>17</sup> diferentes con diferentes dinámicas, que coinciden en su propósito general de mejora (Blalock 1999). Sin embargo, su evolución ha sido independiente y su convivencia un tanto problemática. En

---

<sup>16</sup>En consecuencia, se considera a efectos de la presente tesis que la función de seguimiento (monitoreo) es parte integrante de la gestión por resultados.

<sup>17</sup>Movimientos en el sentido de que los dos han tenido un desarrollo gradual, con investigación teórica y empírica.

este sentido, a pesar de que la evaluación era una práctica establecida durante un período anterior a la emergencia de la gestión por resultados, los roces y tensiones entre los grupos profesionales inmersos en el desarrollo de una u otra fueron frecuentes a lo largo de la década de los años 90.

Ya se ha mencionado al inicio de este apartado que las principales críticas a la evaluación se asentaban en su falta de puntualidad (*timeliness*) y en la escasa influencia que podían tener en la toma de decisiones derivada de la dificultad de aplicar sus recomendaciones. Estas críticas ponen de manifiesto que la evaluación, a pesar de ser una disciplina profesional asentada, era aún entendida como una función externa y puntual, mientras que la gestión por resultados se consideró, desde su concepción, como un instrumento diseñado por y para los equipos responsables de la gestión de los programas y, por tanto, como una práctica integrada en la organización que debía ser liderada internamente.

Por su parte, la comunidad evaluadora, al sentirse directamente cuestionada, no dudó en manifestar su escepticismo con los sistemas de gestión por resultados (Van Thiel y Leew 2002; Greene 1999). Las críticas provenientes del mundo de la evaluación se centraron fundamentalmente en la función de monitoreo como parte de la Gestión por Resultados, y se basaron principalmente en la identificación de tres problemas que afectaban a dicha función (Davies 1999):

- ✓ Falta de calidad técnica: la falta de rigor metodológico en la formulación de indicadores así como de conocimientos evaluativos constituyó una de las críticas más agudas a la función de seguimiento. Esta crítica se debía principalmente a que los indicadores eran formulados por personal interno que no contaba con conocimiento evaluativo.
- ✓ La confusión de objetivos (*goal displacement*): Considerado como el talón de Aquiles del monitoreo, esta confusión de objetivos se produce cuando las instituciones, al ser susceptibles de recibir incentivos para medir e informar sobre ciertos indicadores pero no sobre otros, pierden la visión de cuáles son sus verdaderos objetivos. De esta forma, satisfacer la consecución de los indicadores propuestos pasa a convertirse en el verdadero motor de la gerencia, en vez de ser un medio o instrumento para identificar desviaciones y mejorar los resultados. Este problema puede derivar en situaciones extremas cuando el personal del programa asume, como primera prioridad, el cumplimiento de objetivos en base a los indicadores propuestos, obviando las causas que subyacen en el proceso de la determinación de dichos objetivos y llegando incluso a manipular la información y los indicadores, si esto es preciso, para alcanzar su cumplimiento.
- ✓ Limitaciones en el análisis del contexto y de las causas de los problemas. El 'mantra' o la esencia del monitoreo reflejada en la frase 'lo que se mide, se hace' (*what gets measured gets done*) puede llevar a

dejar fuera de la función de monitoreo a aquellos resultados y efectos del programa que sean difíciles de medir, llegando así a situaciones en las que las raíces de eventuales problemas vinculados al programa en cuestión puedan quedar fuera del radar de la gerencia. Este es un problema frecuente ya que los programas normalmente despliegan una elevada complejidad y muchos de sus resultados no son fáciles de medir a través de indicadores. Las críticas realizadas desde la evaluación (Perrin 1999; Blalock 1999) inciden en este problema llegando a afirmar que, a veces, lo medido no es necesariamente lo más relevante. En este sentido, si la función de monitoreo tiende a centrarse excesivamente en la medición de indicadores, la verdadera complejidad de los programas puede ser peligrosamente ignorada por el proceso de generación de información y, por tanto, por la gerencia.

Al inicio del milenio, el rol de la evaluación con respecto a la gestión por resultados (y, por tanto, a la función de seguimiento) seguía sin estar claramente delimitado en la práctica, no solo en la mayoría de las agencias de cooperación al desarrollo a nivel internacional sino también en los sistemas públicos de los países de la OCDE. El problema fundamental residía en que la relación del monitoreo con la evaluación no estaba claramente definida desde el inicio, lo cual generó diversas tensiones incluso en aquellos ámbitos donde la evaluación era una práctica sólida y bien establecida. Este hecho impulsó la realización de guías por parte de los diferentes organismos, en las que se reflejaban las diferencias entre ambas funciones.

En este proceso de diferenciación de ambas funciones, monitoreo y evaluación, uno de los hallazgos más notables fue que la función de monitoreo, aunque estaba vinculada a la gestión por resultados, era una función que ya existía en la gran mayoría de las organizaciones antes del surgimiento de las teorías vinculadas a la propia gestión por resultados. El monitoreo era una práctica habitual de la gerencia de los programas y estaba centrado fundamentalmente en ofrecer seguimiento a las actividades de los programas en base a los plazos previstos y a los recursos disponibles. Era una actividad inherente a la gerencia. La evaluación, por su parte, cohabitaba sin dificultad con el monitoreo al desenvolverse ésta en un plano más estratégico y analítico, abogando por un enfoque de análisis de resultados y valoración de los efectos producidos.

Sin embargo, el surgimiento de la gestión por resultados ofreció un nuevo carácter a la función de seguimiento tradicional que provocó las tensiones entre ambas disciplinas mencionadas anteriormente<sup>18</sup>. Este nuevo enfoque elevó el perfil del monitoreo ampliando sus funciones analíticas y, por tanto, abarcando parte de las tareas que la práctica evaluativa venía desempeñando.

Con el tiempo, fruto de estas confrontaciones iniciales y del enriquecimiento de los debates posteriores, las posiciones oficiales de los organismos bilaterales y multilaterales que han introducido la gestión por resultados se ha ido modificando con respecto a las funciones de monitoreo y evaluación. En la

---

<sup>18</sup>Tras la incorporación de la Gestión por Resultados, el monitoreo evoluciona hacia lo que se denomina Monitoreo basado en Resultados (*Results Based Monitoring*).

actualidad, existe un acuerdo sobre la necesidad de buscar puntos de encuentro y complementariedades entre ambas. Así, lo confirma la OCDE cuando señala que *“la evaluación y el seguimiento son formas complementarias de los procesos de generación del conocimiento y por tanto ambas deben ser parte integrante del sistema de gestión por resultados”* (OECD 1998, pág. 3). De esta forma, destacar la complementariedad entre ambas funciones representa una práctica común en la literatura evaluadora más actual (Mayne y Zapico-Goñi 1997; Blalock 1999; Rist 2006; Hunter 2006). En este sentido, a continuación se resumen las principales complementariedades entre seguimiento y evaluación (OECD 2010, Rist 2006):

- ✓ El seguimiento se enfoca en medir el grado de cumplimiento de los resultados tal y como estaban planeados (en base a unos indicadores y metas predeterminados). Las evaluaciones pueden tomar los datos obtenidos a través del monitoreo como información de referencia, pero su análisis debe estar más enfocado a identificar las causas que producen y conducen a esos resultados y a identificar los mecanismos que generan los cambios (cómo se ha llegado a estos resultados). En otras palabras, mientras que el seguimiento requiere de un análisis directo de los resultados obtenidos con respecto a la planificación previa, la evaluación generalmente realiza un análisis profundo del contexto y de los mecanismos que explican los resultados logrados. Rist (2006) considera que este hecho justifica la complementariedad secuencial entre seguimiento y evaluación.
- ✓ La función de seguimiento recoge datos a través del monitoreo de indicadores en intervalos específicos de tiempo (trimestral, semestral, anual) en función de la naturaleza del indicador. Sin embargo, las evaluaciones se realizan de forma ocasional (antes, durante y/o después del programa) y la decisión de realizarlas se lleva a cabo generalmente en base a un criterio de selección o ad hoc, en caso de responder a una demanda específica de la gerencia.
- ✓ El seguimiento tiene un carácter general y rutinario. En teoría, la actividad de monitoreo cubre todos los resultados esperados de los programas mientras que la evaluación puede tener un enfoque particular en función del propósito de la misma. Así, por ejemplo, la evaluación puede centrarse en aquellas áreas que han tenido peores resultados o en las que han cosechado resultados inesperados (muy diferentes a lo planeado). En esta misma línea, los informes de monitoreo suelen ser de carácter obligatorio para todos los programas y proyectos mientras que no todos los programas y proyectos son sometidos a evaluaciones.
- ✓ Desde un punto de vista teórico, tanto el seguimiento como la evaluación pueden tratar de responder a preguntas sobre rendimiento. Sin embargo, sus diferencias en cuanto a enfoque y metodología anteriormente descritas, pueden hacer que el monitoreo responda más adecuadamente a unos aspectos particulares y la evaluación a otros

distintos. Por ejemplo, el monitoreo es muy útil para responder a preguntas relativas a la efectividad en cuanto al logro de resultados y a la implementación de actividades. Mientras, la evaluación puede ser más apropiada para responder a preguntas de carácter más cualitativo, como la relevancia de los programas para los colectivos de beneficiarios, la atribución y contribución de los resultados a los objetivos de los programas o la sostenibilidad<sup>19</sup>.

✓ El seguimiento se basa fundamentalmente en el monitoreo de información cuantitativa (incluso los indicadores cualitativos, cuando éstos se elaboran, son generalmente codificados en medidas numéricas), mientras que la evaluación, en muchas ocasiones, dependiendo de las preguntas a responder, utiliza información cualitativa. La información cualitativa de las evaluaciones, a diferencia de la codificada por la función de seguimiento, tiene el fin de explicar y ofrecer mayor detalle sobre áreas concretas de análisis y no busca necesariamente la generalización. Sin embargo, la codificación en valores numéricos de la información cualitativa obtenida en el seguimiento, busca principalmente la generalización de datos.

✓ El seguimiento alerta a la gerencia sobre problemas específicos, pero sus diagnósticos no suelen ir acompañados de recomendaciones. La evaluación en cambio no se limita a mostrar evidencias sobre resultados o dificultades concretas, sino que generalmente acompaña el análisis con conclusiones y recomendaciones sobre cómo enfrentar las debilidades encontradas.

✓ El seguimiento es considerado una labor principalmente de carácter interno llevada a cabo por el personal del programa. La evaluación, por el contrario, tiene un carácter independiente y suele ser desarrollada por personas no vinculadas a la implementación de los proyectos. En cualquier caso, dependiendo de cuál sea el propósito de la evaluación, el criterio de imparcialidad e independencia varía. Así, si el propósito es la rendición de cuentas a organismos externos, la independencia del equipo evaluador es fundamental para dar validez y credibilidad a los hallazgos. Sin embargo, en una evaluación donde el propósito principal es el de aprendizaje, el criterio de independencia pierde cierta relevancia y se acentúa la contribución de la evaluación al empoderamiento, facilitando una amplia participación de actores clave y grupos de interés durante el desempeño de la práctica evaluativa<sup>20</sup>.

---

<sup>19</sup>John Mayne (2001), a través de la teoría de la contribución, afirma que el monitoreo puede dar cierta evidencia, pero no 'pruebas', en la relación de causalidad entre programas y resultados. La función de evaluación, cuando se desempeña a partir de información recogida a partir de las prácticas de monitoreo, se encuentra en una posición más favorable para llevar a cabo análisis más robustos y sustentados de causalidad entre la acción programática y los resultados obtenidos.

<sup>20</sup>Autores como Binnendijk (citado en OECD 2010) y Rist (2006) consideran que la concepción tradicional de que el monitoreo es interno y la evaluación es externa tiende a diluirse. De hecho, muchos organismos impulsan la realización de auto-evaluaciones (internas) como manera de asegurar un mayor

La Tabla 1.2 resume las complementariedades entre monitoreo y evaluación como fuentes de generación de conocimiento en los procesos de toma de decisiones, bien sea para la mejora y el aprendizaje, como para la rendición de cuentas.

**Tabla 1.2. Complementariedades entre Monitoreo y Evaluación**

Monitoreo	Evaluación
- Análisis realizado por personal interno	- Incluye evaluadores independientes
- Niveles de cobertura amplios	- Análisis exhaustivo
- Practica rutinaria	- Ocasional
- Obligatoria	- Opcional, de carácter selectivo
- Frecuentan métodos cuantitativos	- Frecuentan métodos cualitativos
- Responde a la pregunta: ¿Qué resultados se han logrado?	- Responde al cómo y al porqué de los resultados.
- Alerta al personal del programa de problemas	- Proporciona recomendaciones al personal

Fuente: Elaboración propia, adaptada de Binnendijk (citado en OECD 2010, pág. 112)

El monitoreo y la evaluación como fuentes de conocimiento organizacional, también, pueden ser complementarios en relación a los tipos de usos que se deriven de cada uno de ellos. Por ejemplo, puede tener más sentido que de la práctica evaluativa se derive un uso conceptual<sup>21</sup> o incluso el uso de proceso<sup>22</sup> previamente analizados, mientras que es más factible y frecuente que se deriven usos instrumentales de actividades relacionadas con el monitoreo. Igualmente, es frecuente que ambas disciplinas en un marco cooperativo sean capaces de establecer sinergias y complementariedades que mejoren los flujos de información y, por ejemplo, incrementan la capacidad de la práctica evaluativa para producir un uso instrumental. En esta línea, Nielsen y Ejler (2008) han realizado un ejercicio de comparación de los tipos de uso y su factibilidad en actividades de monitoreo y evaluación. Esta contribución queda recogida en la Tabla 1.3.

---

empoderamiento de la gerencia en el análisis de los resultados. De la misma manera, cada vez hay más voces críticas con la aclamada falta de validez de los resultados obtenidos a través del monitoreo interno. A su vez, progresivamente cobra más importancia la tendencia a validar desde un punto de vista externo los resultados obtenidos internamente por las funciones de seguimiento (auditorías, controles de calidad etc.).

<sup>21</sup>Orientado a generar cambios en la forma de entender y concebir cuestiones estratégicas.

<sup>22</sup>Persigue mejoras a nivel cognitivo, de comportamiento o incluso a nivel organizacional a partir de las enseñanzas provenientes del propio proceso evaluativo.

**Tabla 1.3. El tipo de utilización y su factibilidad en Monitoreo y Evaluación**

Tipos de uso	Evaluación	Monitoreo
Uso conceptual	Alta. Puede cuestionar la relevancia del programa.	Baja. Bajas posibilidades de cuestionar la existencia del programa.
Uso instrumental	Media a baja. En ocasiones, es difícil la sincronización de la evaluación con el ciclo de toma de decisiones.	Alta, medidas correctivas pueden ser aplicadas más rápidamente por el personal.
Uso de proceso	Alta, en especial en evaluaciones participativas	Alta, si la información procede de las bases (bottom-up process)

Fuente: Elaboración propia, adaptada de Nielsen y Ejler 2008, pág. 180

Recapitulando, las teorías más recientes confirman que el monitoreo y la evaluación, a pesar de sus diferencias, sirven de manera complementaria a la generación de conocimiento organizacional. En muchos casos, la generación de información a través de indicadores de seguimiento puede facilitar la realización de la función de evaluación, posibilitando una mayor profundización en el análisis de las causas ante la consecución o no de los resultados esperados.

En cuanto a su utilización, se puede afirmar que ambas fuentes (monitoreo y evaluación) presentan diferentes características y, por tanto, pueden ser proclives a diferentes tipos de uso, tal y como se ha expuesto. Una de las mayores limitaciones que se le atribuye tradicionalmente a la evaluación es su dificultad de sincronizar los resultados evaluativos con el ciclo programático y, en consecuencia, su incapacidad para influir de forma puntual en la toma de decisiones en el ámbito de programas. Es difícil concebir un uso puramente instrumental a partir de la práctica evaluativa, mientras que si ésta se apoya en los flujos de información provenientes del seguimiento y monitoreo, su capacidad para resolver el problema previamente planteado se ve incrementada.

La utilización en evaluación ha evolucionado hacia una concepción del uso multidisciplinar donde la temporalidad, la intención y la fuente de información son aspectos a tener en cuenta. Esta nueva concepción enriquece su interpretación teórica, a la vez que supone un mayor desafío desde el punto de vista empírico. A las ya existentes limitaciones para mostrar los factores determinantes del uso más tradicional, se une la dificultad de verificar aspectos no intencionales y de largo plazo, recogidos en los enfoques más recientes de influencia.

Las dificultades de la evaluación para generar cambios en el nivel superior de toma de decisiones, unidas a las dificultades inherentes para medir el grado de utilización de las evaluaciones, han puesto de relevancia la necesidad de

concebir la evaluación como parte integrante de los sistemas de información que nutren de conocimiento a una organización.

En este escenario de convivencia de la evaluación con otras disciplinas generadoras de información valiosa para las organizaciones, emerge un concepto nuevo que consolida la práctica evaluativa dentro de las instituciones. El siguiente capítulo se centra en el concepto y en los elementos vinculados con el desarrollo de capacidad evaluativa y, por ende, en la habilidad de las evaluaciones para generar conocimiento útil y sostenible.

## 2. La capacidad evaluativa de las organizaciones

### 2.1. El concepto de capacidad evaluativa

El hecho de que la práctica evaluativa forme parte integrante del sistema de generación de conocimiento de un organismo determinado y de que sea considerada como un elemento de utilidad en la toma de decisiones depende en gran medida de la capacidad evaluativa de dicha organización. De esta forma, el concepto de capacidad evaluativa hace referencia, en términos generales, a los diferentes esfuerzos desplegados en las organizaciones para integrar la práctica de la evaluación en los procesos de toma de decisiones a niveles gerenciales y estratégicos.

La capacidad evaluativa es, por tanto, un concepto inherente a la práctica de la evaluación, cuyo interés se ha visto incrementado notablemente en los últimos años. En este sentido, la concepción de la función de evaluación como una herramienta para el aprendizaje organizacional ha ido progresivamente consolidando ésta práctica dentro de los organismos y suscitando amplios debates sobre cómo fomentar el establecimiento y el desarrollo de la capacidad evaluativa (Baizerman et al. 2002).

Autores como Mackay (2002), De Lancer Julnes (2006), Mayne y Rist (2006), Nielsen y Ejler (2008) y Rist (2006) coinciden en la necesidad de desarrollar la capacidad evaluativa en las organizaciones con el fin de que el conocimiento generado por el monitoreo<sup>23</sup> y la evaluación sea de utilidad en los procesos de toma de decisiones. En estos trabajos, se considera que la capacidad evaluativa es el vehículo que hace posible transitar desde una concepción de la práctica evaluativa limitada a la elaboración de estudios aislados hacia una visión estratégica donde las evaluaciones se constituyen como ‘torrentes de conocimiento’ y por tanto son susceptibles de ser aplicadas en instancias vinculadas a la gobernabilidad institucional<sup>24</sup>.

En la literatura sobre evaluación encontramos varias definiciones sobre capacidad evaluativa y sobre sus implicaciones en las organizaciones.

Boyle, Lemaire y Rist consideran que el ‘régimen de la evaluación’ se refiere a la creación de las capacidades que permitan el ejercicio adecuado de la práctica evaluativa, incluyendo no solo aspectos organizacionales sino también la institucionalización de la función de evaluación. “*El desarrollo de la capacidad evaluativa se refiere a las actividades impulsadas para poner en marcha dicho régimen*” (Boyle et al. 1999, pág.6)

Una de las tareas inherentes, por tanto, a la capacidad evaluativa es la de garantizar la sostenibilidad de la función de evaluación a través de su institucionalización. De acuerdo a Boyle et al. (1999, pág. 12), “*Conseguir un*

---

<sup>23</sup>La función de monitoreo es entendida en este período como una práctica complementaria a la función de evaluación, superando así las tensiones que surgieron originalmente entre ambas con la aparición de las teorías vinculadas a la gestión por resultados, analizadas en el apartado 1.7.

<sup>24</sup>A pesar de que en la literatura también se aplica el término de capacidad evaluativa a nivel macro (Furubo y Sandhal 2002; Mackay 2002); y micro (Taut 2007), a efectos de esta tesis, la capacidad evaluativa se centra en el nivel organizacional.

*balance equilibrado entre demanda y oferta de trabajo evaluativo supone un ejercicio clave para la institucionalización de la función de evaluación*". Para ello, la evaluación debe pasar de ser una actividad aislada y de corto plazo, a una función integrada en un sistema de conocimiento organizacional estable y sostenible (Schaumburg Muller 1996).

Muchos de los debates sobre capacidad evaluativa se han centrado, por tanto, en identificar aquellos factores que faciliten la integración de las evaluaciones en el conocimiento organizacional. En 2002, Gibbs, Napp, Jolly, Westover y Uhl, identificaron algunos de estos factores y, en concreto, señalaron que, para facilitar la institucionalización de la función evaluativa en un organismo determinado, los recursos disponibles con este fin y la motivación de individuos y grupos clave dentro de la organización en cuestión, son elementos que adquieren una especial relevancia. En consecuencia, estos autores definen la capacidad evaluativa como *"el grado de recursos y motivación en una organización para conducir, analizar y utilizar las evaluaciones"* (Gibbs et al. 2002, pág. 261).

Por su parte, en el año 2002, Stockdill, Baizerman y Compton ofrecieron una definición que ha sido ampliamente utilizada como referencia por la mayoría de los expertos que, más tarde, han profundizado sobre este concepto:

*"La capacidad evaluativa hace referencia a un sistema de acciones intencionales dentro de un organismo, dependientes de un contexto, que incluye procesos y prácticas guiadas para conseguir de manera sostenible que las evaluaciones de programas sean de alta calidad y su utilización apropiada, regular y natural. Capacidad evaluativa es por tanto el trabajo intencional de crear y sostener de manera continua una serie de procesos organizacionales que conducen a una práctica evaluativa de calidad y a un uso rutinario de la misma"* (Stockdill et al., 2002, pág. 8).

Debido a la importancia que esta definición ha tenido en los posteriores modelos teóricos, se realiza a continuación un análisis más detallado, destacando los elementos principales de la misma:

- ✓ Propósito: Los motivos inherentes al establecimiento de la capacidad evaluativa en las organizaciones son principalmente dos (Milstein et al. 2002; Stockdill et al. 2002; Cousins et al. 2004): una mejora de la calidad de la evaluación y un uso sostenible de la misma<sup>25</sup>, ambos orientados hacia el objetivo último de mejorar la gobernabilidad (OECD/DAC 2009). Así, la aplicación de estos propósitos (mejora de la calidad y uso sostenible) en el ámbito organizacional es capaz de fortalecer la capacidad institucional en los siguientes ámbitos (King y Volkov 2005, King 2007): i) diseño, implementación y gestión de evaluaciones efectivas, ii) acceso, desarrollo y utilización de conocimiento evaluativo, y iii) desarrollo de un espíritu de aprendizaje organizacional continuo.

---

<sup>25</sup>La definición abarca el concepto de uso de manera global, donde los diferentes tipos (instrumental, conceptual, de proceso), están incluidos.

✓ Contexto dependiente: La manera de fomentar la capacidad evaluativa depende de la realidad de la organización y de su contexto. Patton (2007) y Taut (2007) coinciden en la complejidad del concepto de capacidad evaluativa y en la necesidad de flexibilizar y adaptar su definición al contexto donde se va a desarrollar. Otras aportaciones también coinciden en aplicar esta flexibilidad a los niveles de intencionalidad y/o a los propósitos de las actividades relacionadas con la capacidad evaluativa (Cousins et al. 2004).

✓ Carácter intencional: La capacidad evaluativa surge a partir de un sistema de acciones intencionales. La definición de capacidad evaluativa incide en el carácter intencionado de las acciones necesarias para ponerla en marcha y para que sea efectiva. Preskill y Boyle, a través de su contribución al concepto de capacidad evaluativa, lo explican así: “*El desarrollo de la capacidad evaluativa incluye el diseño e implementación intencional de estrategias de enseñanza y aprendizaje para asistir a las personas a nivel individual, así como a los colectivos y a las organizaciones en sí mismas en el aprendizaje de lo que constituye una práctica evaluativa efectiva, útil y profesional*” <sup>26</sup>(Preskill y Boyle 2008, pág. 444).

✓ Desarrollo de prácticas y procesos: Para poner en marcha acciones intencionales que deriven en el desarrollo de la capacidad evaluativa en una organización determinada se requiere de un sistema de prácticas y mecanismos. El carácter ‘sistemático’ de estas medidas otorga importancia al rigor procedimental necesario que contribuye a la institucionalización de la función de evaluación. Preskill y Boyle señalan que el objetivo último de la capacidad evaluativa es la práctica evaluativa sostenible, donde las personas de la organización se interesan y preguntan acerca de la recogida y del análisis de datos y usan los hallazgos de la evaluación en la toma de decisiones y en la puesta en marcha de acciones determinadas. Para que la evaluación sea una

---

<sup>26</sup>La literatura especializada recoge tres formas de aprendizaje que, por su importancia, se explican a continuación:

a) el *aprendizaje individual*, el cual se orienta hacia nuevas experiencias de conocimiento y se origina en las personas. Nonaka y Takeuchi (1999) consideran que como resultado de la experiencia particular en contacto con los hechos y con la realidad, se genera un conocimiento personal que se ve fortalecido al ser éste compartido con los demás. Dicho aprendizaje destierra una actitud de aproximación pasiva hacia las cosas;

b) el *aprendizaje de equipo*, el cual apunta hacia la necesidad de crear condiciones y mecanismos para que emerjan grupos específicos orientados al aprendizaje ya que se parte de la idea de que las personas son más inteligentes a nivel colectivo que a nivel individual. Autores como López (2003), a través de estudios empíricos, afirman que el cociente intelectual de individuos ubicados en pequeños grupos heterogéneos y en ambientes de trabajo colaborativos es potencialmente superior al coeficiente intelectual de individuos que operan de forma aislada. Desde esta perspectiva, el aprendizaje se convierte en una actividad social a través de diversas instancias colaborativas que permiten recoger e integrar distintas experiencias en torno a una comunidad donde unos aprenden de otros, y

c) el *aprendizaje organizacional*, observado como actividad puramente social donde el conocimiento se pone en práctica, se critica y se integra en forma de oportunidades para optimizar el aprendizaje y la efectividad en el ámbito laboral. Supone una visión social y activa de las personas, las cuales mientras se encuentran desempeñando las tareas diarias, inician una serie de procesos a través de los cuales exploran, construyen y descubren nuevos conocimientos, en una búsqueda continua de respuestas y soluciones para la mejora del ámbito social en el que se encuentran inmersas (Robbins 1998).

práctica sostenible, los y las participantes en la misma deben ser respaldados en cuanto a su liderazgo, mediante los incentivos y recursos adecuados, de forma que su aprendizaje sobre la evaluación pueda luego transferirse a sus tareas diarias. A su vez, la práctica evaluativa sostenible también requiere del desarrollo de sistemas, procesos, políticas y planes que ayuden a facilitar la incorporación de la evaluación en las instancias donde la organización persigue su misión y sus objetivos estratégicos (Preskill y Boyle 2008).

Desde las diferentes perspectivas teóricas centradas en la relación que existe entre la generación de conocimiento y las evaluaciones, la importancia de la capacidad evaluativa emerge de forma imperante. Los elementos descritos hasta ahora reflejan el giro que se produce en este sentido, un giro que apunta hacia una nueva visión de la práctica evaluativa, donde ésta deja de ser una tarea aislada y se convierte paulatinamente en una función estratégica integrada en las organizaciones. Esta nueva visión de la evaluación trasciende del ámbito de la teoría organizacional y se extiende al marco institucional en el ámbito global. Así, en el área de la cooperación internacional y de la ayuda oficial al desarrollo (AOD) en particular, hay dos acontecimientos que han tenido lugar en la última década y que han sido particularmente significativos en relación a la necesidad de fortalecer la capacidad evaluativa. Ambos eventos por su importancia y su carácter ilustrativo sobre los cambios acaecidos son analizados en el siguiente apartado.

## **2.2. La capacidad evaluativa en el ámbito de la ayuda oficial al desarrollo**

Después de la Segunda Guerra Mundial se produce un fenómeno sin precedentes en el ámbito internacional. Con el fin de evitar conflictos y preservar la paz a nivel mundial, emerge un amplio entramado institucional de carácter multilateral que da lugar al Sistema de las Naciones Unidas (NU) que hoy conocemos. Así, en 1945, 50 países firman la Carta de las Naciones Unidas la cual se constituye en el tratado internacional fundador del Sistema de NU y da pie al surgimiento de un gran número de organismos y agencias especializadas vinculadas al mismo.

Ese mismo año se funda así la Organización de las Naciones Unidas para la Alimentación y la Agricultura (*Food and Agriculture Organization, FAO*) con el objetivo de erradicar el hambre en el mundo. Un año más tarde, se funda la primera agencia especializada, la Organización Internacional del Trabajo (*International Labor Organization, ILO*) a la cual le seguirán otras agencias especializadas como la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (*United Nations Educational, Scientific and Cultural Organization, UNESCO*) en 1946 o la Organización Mundial de la Salud (*World Health Organization, WHO*) en 1948. Más tarde, en 1965 nace el Programa de las Naciones Unidas para el Desarrollo (*United Nations Development Programme, UNDP*) como uno de los mayores esfuerzos internacionales que liderará el impulso global hacia el desarrollo sostenible y la reducción de la pobreza en el mundo.

Éstos son solo algunos ejemplos de agencias y organismos relevantes vinculados al Sistema de NU. Actualmente, bajo el amparo y guía de seis órganos de gobierno donde 193 países miembros<sup>27</sup> se encuentran representados (la Asamblea General, el Consejo de Seguridad, el Consejo Económico y Social, el Consejo de Administración Fiduciaria, la Secretaría y la Corte Internacional de Justicia), el Sistema de NU se compone de un total de 15 agencias especializadas, 11 programas y fondos y 4 organizaciones afines.

En definitiva, el espíritu que subyace en el Sistema de NU, convertido en la red institucional multilateral más relevante de nuestros días, se ve reflejado en el primer artículo de la Carta de las Naciones Unidas:

*“Los propósitos de las Naciones son: (...)1. Mantener la paz y la seguridad internacionales (...)2. Fomentar entre las naciones relaciones de amistad basadas en el respeto al principio de la igualdad de derechos y al de la libre determinación de los pueblos (...) 3. Realizar la cooperación internacional en la solución de problemas internacionales de carácter económico, social, cultural o humanitario, y en el desarrollo y estímulo del respeto a los derechos humanos y a las libertades fundamentales de todos, sin hacer distinción por motivos de raza, sexo, idioma o religión”<sup>28</sup>.*

Así, el elenco de propósitos sobre los que se asienta el Sistema de NU señala a la cooperación internacional como uno de sus elementos fundamentales al proponer explícitamente la aportación de soluciones a los problemas de índole económico y social que acontecen, a nivel global, a través de la cooperación internacional. En este sentido, el Sistema de NU se convierte en el escenario clave elegido por los representantes gubernamentales de los países para reducir la pobreza global y trabajar en aras a un desarrollo sostenible de los pueblos y naciones. A su vez, la Ayuda Oficial al Desarrollo (AOD) emerge como la herramienta principal para lograr dicho cometido.

El rol representativo de la AOD como instrumento para reducir la pobreza y promover el desarrollo económico de los países más desfavorecidos, se confirma a través de la creación de la OCDE en 1961 y, específicamente, del Comité de Asistencia al Desarrollo (*Development Assistance Committee, DAC*), desde donde los países desarrollados coordinan sus estrategias de ayuda y realizan un registro anual de sus contribuciones. Es precisamente este Comité el que acuña por primera vez el término de ‘Ayuda Oficial al Desarrollo’ en 1969 ofreciendo una definición explícita del mismo que despeja cualquier duda sobre la herramienta elegida por los países desarrollados para cooperar en el ámbito internacional.

*“AOD son aquellas transacciones cuyo objetivo principal es promover el desarrollo económico y social de los países en desarrollo a través de donaciones y préstamos que son concesionales en carácter e incluyen un elemento no reembolsable de al menos el 25% (calculado a una tasa de descuento del 10%)” (OCDE 2006, pág. 16).*

---

<sup>27</sup>Ver el sitio <http://www.un.org/en/members/growth.shtml>

<sup>28</sup> Ver <http://www.un.org/en/sections/about-un/funds-programmes-specialized-agencies-and-others/index.html>

Después de más de medio siglo de AOD y de contribuciones a los países en vías de desarrollo, las cuales agregadas superan el trillón de USD, la comunidad internacional reconoce a inicios del milenio que el aparato de cooperación internacional establecido para ayudar a los países más desfavorecidos puede ser sustancialmente mejorado. Algunos de los esfuerzos internacionales más reconocidos en este sentido son las series de foros globales de alto nivel sobre la eficiencia de la ayuda organizados por la OCDE en Roma, París, Accra y Busan en los años 2003, 2005, 2008 y 2011 respectivamente. En dichos eventos, delegados gubernamentales procedentes de países desarrollados y de países en vías de desarrollo, junto con una amplia representación de la sociedad civil, parten del reconocimiento de que la ayuda oficial al desarrollo no estaba produciendo los resultados esperados. Sobre esta base, los países, a través de sus representantes, dan inicio a una serie de esfuerzos, ampliamente reconocidos y sin precedentes, que están orientados a sugerir una serie de medidas específicas que transformen las prácticas hasta ahora aplicadas en una nueva forma de cooperación internacional donde la ayuda al desarrollo pueda cumplir su cometido<sup>29</sup>.

En particular, es en la Declaración de París sobre la Eficacia de la Ayuda al Desarrollo donde se percibe un giro importante en relación a la forma de implementar los programas y proyectos de ayuda al desarrollo así como en relación al papel de la práctica evaluativa en las instituciones multilaterales.

Firmada en Marzo del 2005 por más de 100 países donantes y receptores de ayuda, así como por un número representativo de agencias internacionales y organismos multilaterales, la Declaración de París es un acuerdo internacional que compromete a sus adherentes a mejorar no solo la eficacia y eficiencia sino también la transparencia de las contribuciones realizadas en el marco de la ayuda oficial al desarrollo. Ya, en el Primer Foro de Alto Nivel sobre Armonización que tuvo lugar en Roma en febrero de 2003, los países donantes y receptores identificaron conjuntamente la necesidad de modificar las prácticas tradicionales de cooperación internacional y mejorar sustancialmente la coordinación entre ellos, simplificando procedimientos innecesarios e incluyendo una serie de medidas orientadas a mejorar la eficacia de la ayuda. De forma más específica, la Declaración de París, considerada como un punto de inflexión en el área internacional de la ayuda al desarrollo, incluye los siguientes principios<sup>30</sup>:

- ✓ Apropriación: Los países en desarrollo establecen sus propias prioridades de desarrollo orientadas a fortalecer sus instituciones. Ejercen así un liderazgo efectivo sobre sus propias políticas así como sobre la implementación de la ayuda que reciben.
  
- ✓ Alineamiento: Los donantes alinean su ayuda a las prioridades de los países en desarrollo y hacen el mejor empleo posible de los planes, políticas y sistemas de cada país.

---

<sup>29</sup> Extraído de: <http://www.oecd.org/dac/effectiveness/thehighlevelforaonaideffectivenessahistory.htm>

<sup>30</sup> Declaración de París: <http://www.oecd.org/dataoecd/53/56/34580968.pdf>

- ✓ **Armonización:** Los donantes coordinan entre ellos las diferentes actividades que vienen desempeñando en el área de cooperación internacional y comparten información relevante al respecto para evitar duplicaciones, posibles descoordinaciones así como para reducir la burocracia innecesaria.
  
- ✓ **Gestión orientada a los resultados:** Los países en desarrollo y los países donantes enfocan sus actividades hacia la generación de resultados que puedan ser medidos en base a una serie de marcos de monitoreo y evaluación previamente establecidos.
  
- ✓ **Responsabilidad mutua:** Se establece la premisa de que los países donantes y los países en desarrollo son solidariamente responsables de los resultados que logran, no solo entre sí, sino también ante órganos legislativos y ante la sociedad civil.

**Figura 2.1. Triángulo de Principios de la Declaración de París.**



Fuente: Sanahuja 2007, pág. 95

Tres años más tarde, en Septiembre de 2008, se celebra el Tercer Foro de Alto Nivel, en Accra con el fin de dar un nuevo impulso a los objetivos trazados en la Declaración de París. En Accra, se enfatiza la idea de que para conseguir una apropiación real de las políticas de desarrollo se necesita un diálogo amplio con participación representativa de la sociedad civil, de los órganos legislativos de los países y de las autoridades gubernamentales a nivel local.

Este tercer Foro de Alto Nivel concluyó con lo que se denomina el Programa de Acción de Accra, que persigue dar impulso a los objetivos y principios de la

Declaración de París mediante los siguientes compromisos entre sus partes firmantes<sup>31</sup> (AGCI 2011):

- ✓ Previsibilidad: Los donantes anunciarán a los países con los que colaboran, en la medida de lo posible, las estimaciones sobre sus planes de ayuda para los siguientes períodos (de tres a cinco años).
  
- ✓ Sistemas nacionales: Los países en desarrollo fortalecerán su capacidad de gestión de fondos procedentes de la cooperación internacional e incrementarán la transparencia en relación al manejo de los mismos. Por su parte, los donantes distribuirán sus contribuciones de acuerdo con los sistemas nacionales definidos por los países receptores.
  
- ✓ Condicionalidad: Las condiciones de los donantes para la entrega de la ayuda estarán basadas y alineadas con los objetivos de desarrollo de los propios países receptores.
  
- ✓ Ayuda desligada: La ayuda no estará sujeta a restricciones que impidan a los países en desarrollo comprar los bienes y servicios que necesitan prevaleciendo así el criterio de mejor calidad al menor precio posible.

Como ya se ha mencionado, tanto la Declaración de París como el Programa de Acción de Accra son considerados como un punto de inflexión en la esfera de la cooperación al desarrollo: las autoridades nacionales, jefaturas de agencias y altos cargos oficiales se comprometen, por primera vez, a aunar esfuerzos para armonizar, alinear y gestionar la ayuda al desarrollo en base a la gestión por resultados. De la misma manera, también por primera vez, se incluyen indicadores y metas medibles para valorar la calidad y la efectividad de la ayuda. En este cometido, la transferencia de empoderamiento es clave de tal manera que, conjuntamente, los países socios ejerciten un liderazgo efectivo en sus políticas y estrategias, coordinando las acciones de desarrollo<sup>32</sup> (OCDE 2005).

Los nuevos compromisos adquiridos implican un mayor interés en mejorar la eficacia de la ayuda, a la vez que apuntalan la relevancia del concepto de capacidad evaluativa. La capacidad evaluativa profundiza en la necesidad de adquirir conocimiento especializado, habilidades particulares y actitudes proclives para garantizar que la práctica de la evaluación influya en la toma de decisiones basada en resultados. En palabras de Segone (2007, pág. 27):

*“El principio de empoderamiento significa que los países son responsables de liderar sus propios sistemas de monitoreo y evaluación. Los donantes y las*

---

<sup>31</sup>Programa de Acción de Accra  
<http://www.oecd.org/dataoecd/58/19/41202043.pdf>

<sup>32</sup>Como explica Sanahuja (2007, pág. 94): “La Declaración de París define un nuevo paradigma en la relación entre donantes y receptores, a los que en el espíritu de corresponsabilidad de la Declaración se les denomina países socios”.

*organizaciones internacionales deben apoyar el desarrollo de capacidades evaluativas que aseguren la sostenibilidad de la información generada y su alineamiento con estándares evaluativos. Este apoyo debe tomar en cuenta el valor de la diversidad de los enfoques evaluativos”.*

Posteriormente en 2011, la Alianza de Busan para la Cooperación Eficaz al Desarrollo reitera de manera firme el compromiso de lograr resultados sostenibles. En este sentido, se apela a la realización de mayores esfuerzos para el logro de resultados, y a la necesidad de mejorar tanto la gestión de las actividades de monitoreo y evaluación como la comunicación de los progresos en las metas propuestas (Párrafo 12.b)<sup>33</sup>. La Declaración de Busan reafirma la importancia del empoderamiento y el fortalecimiento de las capacidades nacionales en la gestión de su propio desarrollo.

Más reciente, como resultado de la Tercera Conferencia Internacional sobre Capacidades Nacionales de Evaluación celebrada en 2013 en São Paulo, Brasil, el año 2015 ha sido declarado Año Internacional de la Evaluación. Esta iniciativa parte con el propósito de que la evaluación sea considerada en la agenda de los Objetivos de Desarrollo Sostenible (ODS), y ha sido refrendada por la resolución A/RES/69/237<sup>34</sup> del Segundo Comité de la Asamblea General de NU en Diciembre de 2014. En concreto, la Resolución A/RES/69/237 solicita a cada país miembro que fortalezca su capacidad para la realización de evaluaciones, y que informe de sus progresos al respecto en el año 2016.

### **2.3. Evaluación y capacidad evaluativa**

Tener conciencia de la importancia del desarrollo de las capacidades evaluativas implica disponer de la habilidad de discernir entre capacidad evaluativa y la práctica evaluativa en sí misma.

Bamberger (2000) es uno de los primeros autores en contrastar ambos conceptos (capacidad evaluativa y práctica evaluativa) en el área de la cooperación al desarrollo. Para ello, primero define los programas de desarrollo como aquellos programas con fines de crecimiento económico y mejora social a partir de las contribuciones realizadas por países desarrollados a través de la ayuda oficial al desarrollo o a través de otras organizaciones y asociaciones de la sociedad civil. La práctica de la evaluación se centra en este ámbito sobre programas y proyectos específicos para la mejora de los mismos o puede ir más allá y buscar, a través de la práctica evaluativa, el desarrollo de capacidades evaluativas en los organismos que implementan los programas y proyectos que permitan replicar sus métodos (Bamberger 2000, pág. 96).

Stockdill, Baizerman y Compton (Stockdill et al. 2002), por su parte, profundizan acerca de las diferencias conceptuales entre evaluación y capacidad evaluativa. En su análisis, consideran incluso posible que los/as evaluadores/as puedan ser responsables de instaurar un mayor grado de capacidad evaluativa en las instituciones, aunque siempre destacando que su

---

<sup>33</sup> Busan Partnership for Effective Development Cooperation: Fourth High Level Forum on Aid Effectiveness, Republic of Korea, 29 November – 1 December 2011.

<sup>34</sup> A/RES/69/237 Building capacity for the evaluation of development activities at country level

rol y responsabilidades en este caso sería diferente al que desempeñarían cuando realizan evaluaciones. La tabla 2.1 refleja las diferentes perspectivas existentes entre evaluación y capacidad evaluativa.

**Tabla 2.1. Evaluación y capacidad evaluativa**

	<b>Evaluación</b>	<b>Capacidad Evaluativa</b>
<b>Concepto</b>	Análisis y valoración sistemática del mérito del objeto evaluado en base a una serie de criterios previamente establecidos.	Sistema de acciones intencionadas, procesos y prácticas guiadas para conseguir que las evaluaciones sean de calidad y que su uso apropiado sea una práctica común y rutinaria.
<b>Papel del evaluador/a</b>	Llevar a cabo la evaluación y facilitar su uso a través de tareas relacionadas con la difusión y comunicación de resultados.	Promover la evaluación como una parte necesaria de las tareas diarias de la organización, integrando ésta en la cultura organizacional.

Fuente: Elaboración propia adaptada de Stockdill et al. 2002.

Aunque los métodos existentes para el desarrollo de capacidad evaluativa en un organismo determinado pueden tener similitudes a aquellos relacionados con la propia realización de las evaluaciones (en particular cuando se trata de evaluaciones ajustadas a modelos participativos), ambas acciones representan prácticas diferentes. La capacidad evaluativa se concentra en integrar una serie de prácticas y rutinas dentro de una organización de cara a posibilitar que la práctica evaluativa se convierta en una parte integrante, relevante y estratégica dentro de dicha organización al servicio de aquellas instancias vinculadas a la toma de decisiones.

En este sentido, Cousins et al. 2004 han identificado dos formas de desarrollo de capacidad evaluativa: i) directa, que supone la planificación de actividades específicas para el desarrollo de capacidad evaluativa con independencia de que ocurran dentro o fuera del marco de evaluaciones en curso, y ii)) indirecta que surge del resultado de la participación del personal en las propias evaluaciones puestas en marcha dentro del organismo en cuestión.

En esta aproximación, la forma indirecta presenta reminiscencias de los modelos participativos de evaluación, donde las evaluaciones son realizadas en equipo entre el personal experto en métodos evaluativos y el personal vinculado al programa evaluado (Cousins y Chouinard 2012). Sin embargo, existen algunas diferencias entre esta forma de desarrollo de capacidad evaluativa indirecta y los modelos participativos en evaluación. Así, por ejemplo, estos procesos de desarrollo de capacidad evaluativa, aunque de forma indirecta, son en última instancia procesos orientados a incrementar la capacidad evaluativa en un organismo determinado, lo que implica que existe una posición institucional inclinada a incrementar dicha capacidad evaluativa a

través de una amplia participación en sus evaluaciones. Este análisis no necesariamente se ha tenido que producir en las organizaciones que optan por realizar evaluaciones participativas *per se*, las cuales pueden ser episódicas y originadas por eventos concretos, carentes de estrategia en virtud de la capacidad evaluativa (Preskill y Torres 1999; Rowe y Jacobs 1998; Stockdill et al. 2002).

En cualquier caso, debido a la influencia que han ejercido sobre los modelos de capacidad evaluativa, en el apartado siguiente se analizan en detalle los modelos de evaluación participativos.

#### **2.4. La influencia teórica del enfoque participativo en la capacidad evaluativa**

La evaluación participativa es una estrategia orientada, en última instancia, a crear una cultura organizacional comprometida con el aprendizaje (King 1995, 2007; Sonnichsen 1999; Patton 2008). Sin embargo, esta estrategia participativa no se centra tanto en los métodos a utilizar para lograr sus objetivos, sino en seleccionar a los actores y grupos de interés adecuados que permitan un óptimo desarrollo del proceso evaluativo. A su vez, la evaluación participativa también se centra en identificar las diferentes formas de inclusión y de representación en el proceso evaluativo de los actores y grupos de interés (Greene 2000).

En la literatura evaluadora no existe, hoy en día, una única aproximación a la evaluación participativa siendo posible identificar múltiples enfoques participativos que divergen sustancialmente en cuanto a sus características identificativas se refiere. Esta falta de uniformidad ha sido resaltada por numerosos/as autores/as quienes valoran positivamente la diversidad en este campo y consideran necesario e importante analizar las similitudes y las diferencias que existen entre los diferentes enfoques participativos (Fetterman et al. 2014; Cousins et al. 2013; Miller y Campbell 2006; Rodríguez-Campos y Rincones-Gómez 2013; O' Sullivan 2004).

Entre las similitudes, todos los enfoques participativos coinciden en reafirmar la relevancia del compromiso de los y las participantes en el uso de la evaluación, lo que ha supuesto una sólida base de referencia para los modelos actuales de capacidad evaluativa (Cousins y Whitmore 1998; Fetterman y Wandersman 2005; Love 2006; O'Sullivan 2004; Rodríguez-Campos 2005).

A su vez, la comunidad evaluadora defensora de los enfoques participativos considera que la participación en los procesos evaluativos tiene un efecto positivo que va más del uso de los hallazgos derivados de la evaluación. A través de la participación, las personas aprenden a pensar críticamente y a formular las preguntas adecuadas que les permiten conseguir la información que necesitan en cada momento, así como la forma adecuada de interpretar los datos, aprendiendo, en definitiva, a emplear el razonamiento lógico en su rutina diaria. Por ello, este proceso de aprendizaje, que se inicia a partir de la evaluación, desarrolla las capacidades de quienes participan en la práctica evaluativa permitiéndoles así afrontar sus tareas diarias de forma adecuada y

sostenible. Horton et al. (2003 pág. 63) afirman: “*El aprendizaje a través de la práctica es fundamental para el desarrollo de capacidades. Por tanto, los gestores que busquen desarrollar dichas capacidades en las organizaciones deben facilitar un entorno abierto a la reflexión, a la crítica y la mejora*”.

Dado que los enfoques participativos han tenido mucha relevancia en la evaluación desde la década de los años 80, es posible hoy encontrar evidencias empíricas sobre sus efectos en la consolidación de una mayor capacidad evaluativa en un organismo determinado, y, en definitiva, en una mayor capacidad para lograr un escenario afín al aprendizaje. Existen así numerosos casos y estudios de evaluación participativa donde se constatan sus efectos positivos para el desarrollo de capacidades evaluativas (Alcocer et al. 1997; Tandon 1981, 2008; Feuerstein 1986).

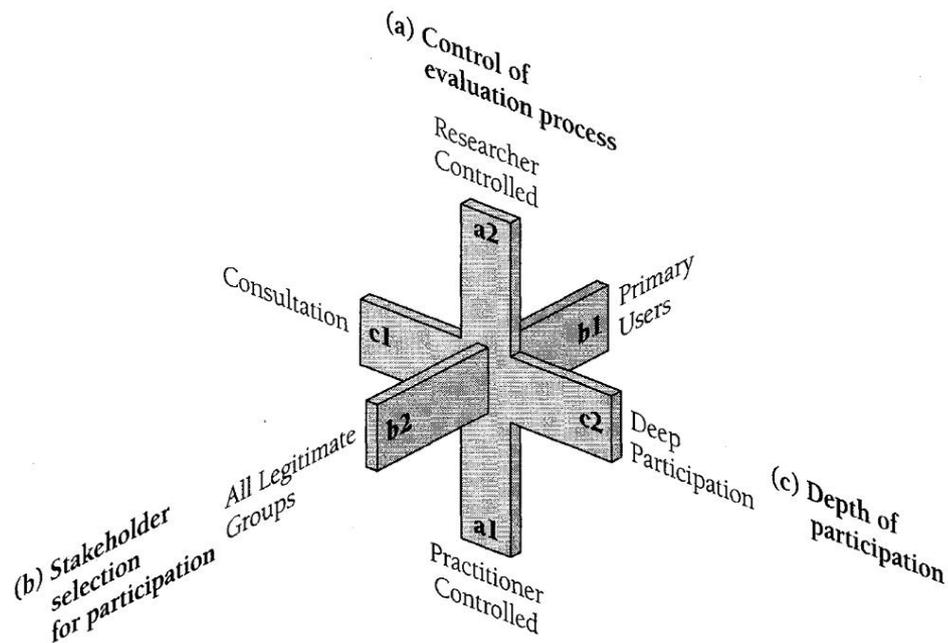
En cuanto a las diferencias entre las distintas formas de desarrollar evaluaciones participativas, muchas de las contribuciones realizadas, hasta el momento, se han centrado en la elaboración de modelos de participación que faciliten su análisis, ya que en este campo se emplea frecuentemente una terminología difusa que no facilita la diferenciación entre las distintas prácticas (evaluación participativa, colaborativa, de empoderamiento, etc.). Este es el caso de los trabajos realizados por Cousins y Whitmore, quienes en 1998 propusieron un modelo para identificar el tipo de participación en las evaluaciones a través de tres dimensiones fundamentales<sup>35</sup>:

- ✓ El control del proceso de la evaluación: esta dimensión abarca desde aquellas evaluaciones donde el/la evaluador/a toma todas las decisiones referentes al proceso de la evaluación, hasta evaluaciones donde el control de las decisiones está totalmente en manos del personal del programa evaluado.
  
- ✓ Selección de los actores clave a los que hacer partícipes: esta dimensión incluye desde evaluaciones donde la participación se limita únicamente a los usuarios primarios y directos, hasta las evaluaciones donde la participación incluye a todos los posibles grupos de interés.
  
- ✓ Grado de participación: esta dimensión diferencia a las evaluaciones en las que se producen niveles mínimos de colaboración (se trata más bien de prácticas consultivas, donde los y las participantes no asumen ninguna responsabilidad en el proceso de evaluación y no tienen un papel destacado en la toma de decisiones), de aquellas evaluaciones en las que se produce un nivel de colaboración intensiva, donde los y las participantes se involucran en todas las etapas del proceso evaluativo y poseen, a su vez, capacidad de decisión asumiendo responsabilidades formales.

---

<sup>35</sup>El modelo fue reformulado más tarde con la incorporación de 2 dimensiones adicionales (Weaver y Cousins 2004). Sin embargo, este cambio ha sido objeto de algunas críticas (Daigneault y Jacob 2009). En consecuencia, la mayoría de la literatura existente se refiere al modelo original de 1998 que, por tanto, es el que se ha seguido en esta tesis.

**Figura 2.2. Dimensiones de los enfoques participativos**



Fuente: Cousins y Whitmore 1998, pág. 11

La gran aportación de este modelo es que facilita el análisis de cualquier tipo de enfoque participativo a través de tres preguntas básicas: ¿Quién toma las decisiones? ¿Quién participa? y ¿Cuánto participa? Además, a pesar de que generalmente una pregunta conduce a la otra, el modelo permite considerar las tres como dimensiones independientes.

Teniendo como referencia este modelo de análisis, a continuación se describen y se analizan los tipos de evaluación participativa más representativos de acuerdo a la literatura existente en la materia:

- ✓ Evaluación participativa práctica o transformativa.
- ✓ Evaluación centrada en la utilización.
- ✓ Evaluación empoderativa.
- ✓ Evaluación colaborativa.

#### **2.4.1 La evaluación participativa práctica y la evaluación participativa transformativa**

Según Cousins y Whitmore (1998, pág. 5), en una evaluación participativa “los investigadores, facilitadores o evaluadores profesionales colaboran en alguna

*medida con individuos, grupos o comunidades que tienen algún interés con el programa, proyecto u organización evaluada*". Esta definición, incluye las personas que Mark y Shotland en 1985 identificaron como 'actores clave', refiriéndose así a aquellos individuos que poseen un cierto grado de interés en la evaluación. Dentro de este grupo, pueden ser incluidos donantes, responsables de la gestión, colectivos beneficiarios, usuarios de la información, etc.

Los autores que han desarrollado este enfoque (Cousins 2004; Cousins y Earl 1995; King 1998) diferencian dos tipos de evaluación participativa: la evaluación práctica y la evaluación transformativa. Ambas evaluaciones tienen en común que los actores clave participan en la evaluación y su principal diferencia reside en el motivo principal o propósito de la evaluación.

**La evaluación participativa práctica** tiene su origen en Estados Unidos y Canadá y su propósito principal es la identificación y solución de problemas específicos para la mejora del programa evaluado. La premisa fundamental sobre la que se asientan este tipo de evaluaciones es que la participación de los actores clave en la evaluación fortalece la relevancia de la información generada y, con ello, su uso en la toma de decisiones. La evaluación práctica puede perseguir varios usos como el instrumental, el conceptual y el uso simbólico. Este tipo de evaluaciones forma parte de muchos de los estudios evaluativos de la década de los años 80, período caracterizado, a su vez, por una tendencia creciente hacia esquemas de responsabilidad compartida entre las personas a cargo de la evaluación y las evaluadas, así como hacia la necesidad de extender el concepto del uso de la evaluación instrumental a otras formas de usos orientadas al aprendizaje organizacional (Cousins y Earl 1995; Jenlink 1994; Owen y Lambert 1995; Ayers 1987; King 1995). Dichos estudios incluían experiencias prácticas de casos particulares donde la participación de actores clave en el proceso evaluativo promovía la utilización de la evaluación, siempre y cuando se dieran una serie de condiciones como la calidad técnica de la misma (Greene et al. 1998; Cousins 1996).

**La evaluación participativa transformativa:** Vinculada a la esfera política, este tipo de evaluaciones persigue lograr un mayor grado de democratización en los países donde se desarrolla, de cara a producir cambios sociales estructurales. Sus orígenes ideológicos son muy diferentes a los de la evaluación participativa práctica. Cousins y Whitmore subrayan que la evaluación participativa transformativa surge en los países en desarrollo, principalmente en América Latina, como una reacción a los modelos positivistas de investigación que eran concebidos como modelos explotadores y aislados de los problemas urgentes de índole económico y social en que sus países se encontraban inmersos (Cousins y Whitmore 1998). Estudios de referencia en la línea descrita han sido elaborados por Fals-Borda (1980), Tandon (1981) y Kassam y Mustafa (1982). Este tipo de evaluación busca actuar como un revulsivo a la situación de poder y control establecida, por lo que, tradicionalmente, la evaluación participativa transformativa ha estado ligada a programas dentro de la esfera de la cooperación internacional al desarrollo, así como a movimientos relacionados con el desarrollo local y de igualdad de género. Las raíces filosóficas de este tipo de práctica evaluativa se

remontan a las tesis de los filósofos alemanes del siglo XIX Karl Marx y Friedrich Engels como padres de la teoría marxista y de la lucha de clases. A su vez, estas ideas se ven reflejadas en el trabajo del brasileño Paulo Freire centrado en la pedagogía crítica (1970) o incluso en las ideas de Jürgen Habermas a través de su influyente teoría de la acción comunicativa. A continuación se destacan varios aspectos clave de la evaluación participativa transformativa:

- ✓ El primero hace referencia a quién crea y controla la generación de conocimiento. En este sentido, el propósito de este tipo de práctica evaluativa es el de empoderar a las personas a través de su participación directa en el proceso de desarrollo de su propio conocimiento (concientización), de tal manera que tengan un mayor entendimiento sobre la relación existente entre conocimiento, poder y control (Fals-Borda y Anisur-Rahman 1991; Tandon 1981). De acuerdo a esta tesis, el conocimiento 'popular' pasa a ser tan útil y aceptado como el conocimiento científico.
- ✓ El segundo aspecto se centra sobre el propio proceso evaluativo, es decir, sobre la forma en la que se produce la evaluación. Ante el interrogante de cuál es la distancia entre evaluador/a y evaluado/a en relación al proceso evaluativo, la respuesta que, desde este tipo de evaluaciones, se ofrecería es que dicha distancia es inexistente, ya que ambos grupos trabajan colectivamente y el diálogo entre ellos es algo rutinario que favorece el entendimiento y respeto mutuo.
- ✓ El tercer aspecto crucial de la evaluación participativa transformativa es su tendencia a favorecer la reflexión crítica. A través de este tipo de evaluaciones se promueve que los y las participantes cuestionen y pongan en entredicho asunciones preconcebidas, así como aquellas normas y protocolos que subyacen de situaciones y prácticas consolidadas en el momento de la evaluación (Comstock y Fox 1993).

Como explican Cousins y Whitmore (1998), bajo el enfoque de las evaluaciones participativas transformativas, la práctica evaluativa es utilizada para transformar las relaciones existentes de poder y promover el cambio y la acción social. La evaluación es concebida como un proceso donde el involucramiento de los y las participantes menos 'poderosos' en la investigación, reflexión, negociación y toma de decisiones, hace que las dinámicas de poder se vean modificadas. En cualquier caso, aunque ambas formas de evaluación - participativa práctica y participativa transformativa - presentan diferencias, también tienen importantes puntos en común. En este sentido, el conocimiento evaluativo adquirido a través de un modelo de evaluación participativa práctica puede llevar al empoderamiento del grupo de gestión dentro de una organización y generar transformaciones de tipo social. De la misma manera, una evaluación transformativa que ha logrado que ciertos grupos menos poderosos aumenten su capacidad de toma de decisiones en un programa, puede también generar un conocimiento práctico y positivo para el

desarrollo del programa específico y su implementación. Por ello, se puede concluir que a pesar de tener orígenes y propósitos principales diferentes, ambas subcategorías de las evaluaciones participativas son altamente complementarias.

A modo de conclusión, si analizamos el tipo de enfoque de evaluación participativa práctica y de la evaluación participativa transformativa dentro del modelo propuesto por Cousins y Whitmore, podemos destacar lo siguiente:

La evaluación participativa práctica: en relación al control de las decisiones (primera dimensión), se trata de un control compartido, donde los actores clave comparten la toma de decisiones con el equipo evaluador. En cuanto a la diversidad de los actores clave (segunda dimensión), nos encontramos en un tipo de evaluación que se enfoca principalmente en los equipos responsables de la gestión del programa y donde otros actores (beneficiarios/as finales, otras organizaciones, donantes) tienen una presencia limitada. Con referencia al grado de participación (tercera dimensión), ésta se encuentra en su nivel máximo de manera consistente a lo largo de todo el proceso.

La evaluación participativa transformativa: el control de las decisiones es balanceado, aunque ligeramente inclinado en favor del evaluador/a que, a través de este control, es capaz de empoderar a los y las participantes<sup>36</sup>. La diversidad de los y las participantes es amplia y facilita que los grupos que tienen menos poder (principalmente beneficiarios/as finales) formen parte del ejercicio evaluativo. En cuanto al grado de participación, el enfoque transformativo es flexible y balanceado aunque más inclinado a favorecer una participación extensiva.

#### **2.4.2 La evaluación centrada en la utilización**

La evaluación basada en la utilización, desarrollada principalmente por Michael Q. Patton aboga por valorar la práctica evaluativa en base a su utilidad y uso<sup>37</sup>. Esto implica que el diseño de la evaluación debe prestar una especial atención a la forma en que cada elemento del proceso evaluativo va a afectar a su utilización. Este enfoque destaca así la importancia que tiene que los usuarios intencionados de la evaluación apliquen los hallazgos de la evaluación a partir de su experiencia y participación en el proceso evaluativo como parte de su trabajo diario. Por esta razón, el foco de atención de la evaluación centrada en la utilización es, precisamente, lograr el uso intencionado por los usuarios intencionados. La característica principal de este tipo de evaluaciones participativas es, por tanto, que las evaluaciones se realicen conjuntamente con usuarios clave previamente seleccionados para usos específicos, también, previamente definidos.

---

<sup>36</sup>Este rol del evaluador/a como sujeto capaz de empoderar ha sido objeto de debate y de diferentes opiniones en la literatura evaluadora. Cousins et al. 2013 consideran que el/la evaluador/a tiene la capacidad y responsabilidad de empoderar. Para Fetterman, esto es algo imposible. Nadie es capaz de empoderarse salvo uno mismo (Fetterman y Wandersman 2007).

<sup>37</sup> También llamada 'enfoque centrado en el uso'.

Para poner en marcha el papel protagonista de los usuarios intencionados en este tipo de evaluaciones, la clave reside en otorgar a dichos usuarios la capacidad de decidir en todos los aspectos relevantes de la evaluación. Por eso, de acuerdo a este enfoque, es fundamental identificar adecuadamente quiénes son los usuarios intencionados. Para ello, en el proceso de selección se parte de un abanico amplio de posibles usuarios, abanico que se va reduciendo hasta llegar a un grupo específico compuesto por usuarios que cumplen con una serie de criterios clave vinculados a la utilización intencionada que va a tener la evaluación correspondiente.

Según este enfoque, el proceso de identificación de usuarios difiere de la identificación de las denominadas 'audiencias' de la evaluación en general. La identificación de usuarios es más precisa y exhaustiva e incluye el análisis de una amplia gama de factores personales con el objetivo de seleccionar personas que tengan un verdadero interés por la evaluación y por los potenciales hallazgos que se generen a partir de ella. Según Patton, la presencia de personas interesadas y activas en el proceso evaluativo es un elemento crucial que potencia su uso (Patton 2008). Así, los análisis sobre el 'factor personal' de los/as candidatos/as incluye muchos aspectos relativos a estas personas como su liderazgo, interés, entusiasmo, determinación y compromiso.

Este énfasis por la identificación del 'factor personal' como elemento clave para la utilización de la evaluación es algo ya identificado por otros diferentes autores/as (Cronbach et al. 1980; Cousins y Leithwood 1986; Preskill y Caracelli 1997). Así, por ejemplo, Cronbach et al. han llegado a afirmar: *"No hay nada que marque más las diferencias en el uso de la evaluación que el factor personal- que es el interés en aprender de la evaluación y el deseo del evaluador en obtener la atención sobre lo que él conoce"* (Cronbach et al. 1980, pág. 6).

El proceso de las evaluaciones centradas en la utilización está compuesto por varias etapas:

1. Identificación de los usuarios intencionados de la evaluación. Estos usuarios son puestos en contacto entre ellos y, en base a un esquema de trabajo predefinido, colaboran estrechamente con el equipo de evaluación y participan junto con éste en la toma de decisiones de la evaluación en general.
2. El personal encargado de la evaluación y los usuarios intencionados definen el tipo de evaluación que desean (por ejemplo, evaluación formativa o sumativa, etc.). Este enfoque servirá para identificar las preguntas de la evaluación. De forma interactiva, el equipo evaluador facilitará el proceso de toma de decisión de los usuarios intencionados.
3. Tercero, los usuarios intencionados adoptan las decisiones oportunas relacionadas con el tipo de técnicas de recogida de información, indicadores y métodos de análisis. Para poder tomar este

tipo de decisiones, los usuarios deben entender las ventajas e inconvenientes de cada uno de los métodos por lo que el/la evaluador/a asume un papel fundamental en esta etapa, facilitando y clarificando aspectos fundamentales relacionados con lo anterior. Las opciones metodológicas son amplias e incluyen métodos de análisis cualitativos así como cuantitativos, basculando, a su vez, entre el mayor o menor grado de énfasis que quiera otorgarse a la realización de generalizaciones. En cualquier caso, la pregunta clave a responder se centra de nuevo la utilidad: ¿serán útiles los métodos seleccionados y cómo contribuirán a que se usen los resultados que producen?

4. Cuarto, una vez que los datos sean recogidos y organizados, los usuarios intencionados participan activamente en su análisis incluyendo la interpretación de los resultados y la generación de recomendaciones. El personal profesional responsable de la evaluación, de nuevo, tiene un papel clave en la facilitación y clarificación de datos y en cómo realizar recomendaciones que puedan ser más tarde utilizadas.

5. Finalmente, se cubren los aspectos más relacionados con la disseminación del informe de evaluación en caso de que éstos no hayan sido tratados en las fases iniciales de diseño. En cualquier caso, los y las defensores de este enfoque explícitamente subrayan la no necesidad de entender el proceso como una secuencia lineal de las diferentes fases sino que, por el contrario, es posible alternar entre ellas si el proceso natural de evolución de la evaluación en cuestión así lo precisa.

Como se puede observar en las fases descritas, dentro de este enfoque de evaluación, el uso del proceso es un uso contemplado y buscado, es decir, el uso del proceso es intencional. Tal y como describe Patton, el proceso de involucrar activamente a los usuarios intencionados incrementa el impacto de la evaluación. Incrementar el grado de aprendizaje a partir de promover en un organismo determinado una actitud positiva ante el propio aprendizaje pueden ser algo intencional y buscado a propósito. En este sentido, el uso no es algo tratado de manera espontánea o informal, sino que es algo planeado e intencionado desde el inicio del proceso evaluativo (Patton 2008).

En suma, los principios base de la evaluación centrada en la utilización son los siguientes:

✓ El compromiso de los usuarios intencionados hacia el uso intencionado es una premisa básica en todo el proceso de evaluación. En todas las decisiones a tomar con respecto al enfoque, diseño, métodos y análisis, la pregunta clave a realizar a los y las participantes es: ¿qué efecto va a tener este hecho en la utilización intencionada de la evaluación?

- ✓ El concepto de uso forma parte de principio a fin de la evaluación y no es algo que surge al final del proceso. De hecho, al final de la evaluación el uso potencial de la misma ya ha sido meticulosamente analizado y promovido.
  
- ✓ El factor personal contribuye significativamente al uso. Este factor parte de la premisa de que el compromiso e interés personal favorece la utilización. Por eso, las evaluaciones deben ser específicamente orientadas a satisfacer las necesidades de los usuarios intencionados.
  
- ✓ El análisis de los actores clave es un paso fundamental en la identificación de usuarios intencionados. El personal, la gerencia, los donantes, y los representantes de las comunidades beneficiarias, por ejemplo, pueden estar entre los usuarios intencionados. Los principios éticos y la sensibilidad política son aspectos clave a considerar en la identificación de estos usuarios intencionados.
  
- ✓ El enfoque es importante para maximizar la efectividad de la evaluación. Los recursos y el tiempo para hacer una evaluación son limitados, y la evaluación basada en la utilización parte de la premisa de que una evaluación no puede responder a todos los interrogantes, ni satisfacer a todo tipo de audiencias. Seleccionar usuarios intencionados para usos intencionados es un hecho que pretende enfocar y, así, limitar el campo de análisis de la evaluación de cara a maximizar su efectividad.
  
- ✓ Las evaluaciones útiles son aquellas diseñadas y adaptadas a un contexto específico. Es importante tener en cuenta en el diseño de una evaluación basada en el uso que los factores situacionales pueden afectar el propio uso de la evaluación. Estos factores incluyen características de la organización, la credibilidad del evaluador/a, consideraciones políticas, limitaciones presupuestarias, etc. En este tipo de evaluaciones, el proceso de la evaluación es activo-reactivo y adaptable y, por tanto, se trabaja en constante coordinación con los usuarios intencionados para valorar el grado en que los factores situacionales pueden afectar el uso de la evaluación en cuestión.
  
- ✓ El compromiso de los usuarios intencionados puede ser fortalecido a través de su participación en la toma de decisiones sobre la evaluación. Este mayor grado de inclusión aumenta la relevancia, el entendimiento y el empoderamiento de la evaluación, factores fundamentales para facilitar su uso.
  
- ✓ La implicación de los usuarios intencionados afectará la calidad y la utilidad de la evaluación. A pesar de que, en principio, puedan existir ciertas reticencias en la comunidad evaluadora a que personas no

expertas en la práctica evaluativa tomen decisiones con respecto a cuestiones metodológicas en el ámbito de la evaluación por temores a sufrir una potencial disminución del rigor metodológico, en este tipo de evaluaciones se considera que la participación plena es el camino para lograr proveer de información útil y precisa a las instancias donde se produce la toma de decisión. De esta forma, los/as evaluadores/as deben facilitar la comprensión de aspectos metodológicos, dotando a los propios usuarios intencionados de la capacidad de valorar y juzgar por sí mismos las diferentes opciones posibles. Puesto que los y las profesionales de la evaluación conocen que su credibilidad e integridad siempre puede estar en entredicho, es importante mantener una actitud activa-reactiva-adaptable. Es decir, el trabajo debe comenzar presentando a los usuarios intencionados cuál es la mejor ruta a seguir en términos de enfoque y método (activa). A su vez, se debe mantener una actitud reactiva y capaces de escuchar y absorber diferentes opiniones e intereses. Por último, el trabajo ha de adaptarse a su entorno sin abandonar, en ningún momento, un escrupuloso respeto por los estándares de calidad y profesionalidad que deben regir la práctica evaluativa. Así, los y las profesionales comprometidos con el uso tienen la responsabilidad de capacitar a los usuarios en los procesos de la evaluación y en los usos de la información generada, fortaleciendo así la capacidad evaluativa hacia el uso.

- ✓ El uso es algo diferente a la preparación del informe de evaluación y su diseminación posterior. El informe final y su difusión pueden ser herramientas importantes para facilitar el uso pero no pueden ser considerados bajo este enfoque como usos intencionados en sí mismos.
- ✓ Tomar en consideración el uso implica tener en cuenta aspectos financieros y temporales. Los beneficios frente a los costes están precisamente representados por un mayor uso de la evaluación. Así, los costes deben ser incorporados, desde el principio, como parte del presupuesto de la evaluación de tal manera que se pueda realizar un seguimiento financiero al grado de utilización.

El enfoque centrado en la utilización, también, vincula la participación y la colaboración con el incremento de la capacidad evaluativa. De hecho, una de las razones de la importancia de la intencionalidad en la participación de los usuarios es la de facilitar un compromiso más a largo plazo en el empleo del pensamiento y de la lógica de la evaluación, desarrollando así una cultura de aprendizaje en el propio programa y en la organización en cuestión. En cualquier caso, el grado de importancia que se le otorgue a este uso de la evaluación es algo que debe ser negociado con los usuarios intencionados.

Este enfoque parte del hecho de que, al igual que el personal de un organismo necesita práctica y experiencia para aprender a realizar evaluaciones, los programas y las organizaciones necesitan, a su vez, práctica y experiencia para incorporar la función de evaluación en los sistemas de aprendizaje organizacional. Es, por esto, que cada vez el campo de la evaluación presta

más atención al desarrollo de la capacidad evaluativa en las organizaciones (Kuzmin 2005).

Así, la evaluación centrada en la utilización considera el desarrollo de la capacidad evaluativa como un aspecto importante para la facilitación del uso de las evaluaciones. Y, de nuevo, considera que su tratamiento debe ser adecuadamente planeado y gestionado, es decir, debe ser intencional, puesto que las actitudes positivas hacia la evaluación y el aprendizaje no tienen por qué considerarse garantizadas. La intencionalidad es un aspecto básico de este enfoque no solo para incrementar el uso de las evaluaciones sino para desarrollar la capacidad evaluativa en las organizaciones.

En cuanto a su localización dentro del modelo de Cousins y Whitmore, el enfoque centrado en la utilización se caracteriza por un control compartido de las decisiones entre los y las evaluadores/as y participantes (primera dimensión); por una participación que se reduce solo a los usuarios intencionados (segunda dimensión); y por un tipo de participación extensivo y regular a lo largo de todo el proceso (tercera dimensión).

Finalmente, de este enfoque se deriva que la apertura a la práctica evaluativa se produce cuando las organizaciones tienen experiencias positivas cercanas con las evaluaciones, experiencias que derivan en el desarrollo de capacidades específicas que permiten absorber mejor las evaluaciones y considerarlas herramientas de utilidad. Construir la capacidad evaluativa es fundamental e incluye muchos elementos, que se analizarán en el apartado 2.5, como la capacitación del personal en la lógica evaluativa, el desarrollo de procesos específicos para integrar la evaluación en el ciclo de los programas y proyectos; y, en definitiva, la conexión de la práctica evaluativa con el aprendizaje organizacional (Sonnichsen 1999; Preskill y Torres 1999).

### **2.4.3 La Evaluación Empoderativa**

Desarrollada por Fetterman, la evaluación empoderativa<sup>38</sup> utiliza los conceptos evaluativos, sus técnicas y los hallazgos de las evaluaciones con el fin de promover la mejora del objeto evaluado (Fetterman 2007). Este tipo de evaluaciones persiguen incrementar las probabilidades de éxito de un programa a través de la provisión de herramientas a los actores clave que les permitan valorar la planificación, implementación y auto-evaluación de sus programas, así como para integrar la evaluación en su planificación y gestión habitual (Fetterman y Wandersman 2005).

Al igual que la evaluación centrada en la utilización, en la evaluación empoderativa el uso del proceso es el tipo de uso más importante (Fetterman 2001; Fetterman y Wandersman 2005). El modelo se basa en una serie de principios básicos que se describen a continuación (Fetterman y Wandersman 2005):

---

<sup>38</sup> También llamada 'enfoque orientado al empoderamiento'.

- ✓ Mejora: la evaluación empoderativa no es un experimento neutral sino que tiene una finalidad clara que consiste en perseguir que la evaluación se utilice en la mayor medida de lo posible y contribuya de manera sustancial a que las personas puedan cumplir sus objetivos.
  
- ✓ Empoderamiento de la comunidad: el sentido de empoderamiento (*ownership*) sobre la evaluación es ampliamente aplicado en la esfera de la cooperación internacional al desarrollo y parte de la idea fundamental de que las comunidades receptoras de la ayuda y en las que se implementan los distintos programas y proyectos, poseen el derecho elemental de ser capaces de adoptar decisiones sobre acciones que afecten a sus vidas.
  
- ✓ Inclusión: la representación de participantes en este tipo de evaluaciones es amplia y diversa, incluyendo a los órganos de gestión del programa, al equipo de evaluación, e incluso a los donantes. La idea que subyace en esta práctica inclusiva es que la diversidad suma y enriquece.
  
- ✓ Participación democrática: la evaluación empoderativa confía en la capacidad de las personas de valorar y emprender las acciones correspondientes, si se producen condiciones adecuadas en un ambiente de consenso, donde la participación es libre y los procesos transparentes.
  
- ✓ Justicia social: la evaluación empoderativa se sustenta sobre valores morales de justicia, persiguiendo compromisos firmes con formas de distribución equitativas así como con esquemas sociales de igualdad de oportunidades y obligaciones. Este tipo de evaluaciones parten del reconocimiento de que existen inequidades sociales básicas en la sociedad de índole moral y las evaluaciones no puede estar deslindadas de dichas inquietudes. De esta manera, la práctica evaluativa es entendida, en última instancia, como un vehículo para mejorar no sólo los programas sino también su impacto en las condiciones sociales bajo un marco de justicia social.
  
- ✓ Conocimiento comunitario: la evaluación empoderativa respeta y valora el conocimiento (*know how*) comunitario, considerado, en todo momento, como parte fundamental de la práctica evaluativa.
  
- ✓ Estrategias basadas en evidencia: a partir del uso de teorías educativas donde el conocimiento básico basado en evidencias empíricas juega un papel prioritario, la evaluación empoderativa incita a las comunidades, donde se desarrollan los programas y proyectos, a adaptar dicho conocimiento básico a las prácticas establecidas en aras a mejorarlas.

- ✓ Desarrollo de capacidades: la habilidad adquirida por los actores clave que participan en la realización de las evaluaciones empoderativas es uno de los pilares básicos sobre los que se asienta la estrategia de dichas evaluaciones para mejorar la planificación e implementación de los programas.
  
- ✓ Aprendizaje organizacional: la utilidad de las evaluaciones empoderativas reside en su capacidad de informar y educar a los equipos de gerencia dentro de un organismo determinado para mejorar los procesos de toma de decisión, proveyéndoles de información estratégica fundamentada a partir de evidencias y donde el papel en su preparación de todas las partes implicadas (comunidades) es crucial.
  
- ✓ Rendición de cuentas: la evaluación empoderativa se produce en un marco de políticas y estándares ya establecidos, donde es importante entender la medida en que el programa ha cumplido objetivos así como la forma en que los actores clave han sido capaces de asumir sus responsabilidades para maximizar resultados.

Según Fetterman (2005), estos diez principios son los que guían la evaluación empoderativa desde su propia conceptualización hasta su desarrollo: *"los principios de la evaluación empoderativa son las lentes que ayudan a enfocar una evaluación"* (Fetterman 2005, pág. 2). Entre ellos se destaca el principio de inclusión el cual sugiere que se trate de buscar la representación más amplia posible de actores clave a pesar de posibles restricciones en cuanto a tiempo y recursos. A su vez, el principio de desarrollo de capacidades es también de gran relevancia para este tipo de evaluaciones. Bajo este principio, se debe otorgar a las comunidades un papel representativo en el proceso de la evaluación como, por ejemplo, brindándoles la oportunidad de recoger sus propios datos, a pesar de que en ocasiones pueda sea más fácil y rápido que estas tareas sean asumidas por los y las profesionales de la evaluación. El principio de rendición de cuentas recuerda, a su vez, a la comunidad que son los actores responsables ante requerimientos externos.

Existen otros conceptos que son clave en la evaluación empoderativa tales como la figura del amigo/a crítico/a, la cultura de evidencia, los ciclos de reflexión-acción, la idea de comunidad de aprendices y las denominadas actitudes reflexivas<sup>39</sup>. A continuación, se definen brevemente cada uno de estos conceptos:

- ✓ La figura de amigo/a crítico/a se emplea en este tipo de evaluaciones para reflejar mejor el papel del evaluador/a, una persona profesional que, aunque comprometida con el programa evaluado, es a

---

<sup>39</sup> Estos conceptos están influenciados por las teorías de desarrollo organizacional y transformativas, incluyendo a Argyris y Schon (1978), Senge (1990), y Preskill y Torres (1999).

la vez crítica y analítica con el mismo<sup>40</sup>. El/la evaluador/a plantea de forma constante interrogantes cuyo análisis permita garantizar el rigor y transparencia necesarios en todo el proceso evaluativo, ya que su objetivo final es que el programa sea más efectivo y cumpla con sus objetivos. En la evaluación empoderativa es, por tanto, un amigo/a crítico/a que ayuda a facilitar el proceso pero no lo controla, limitándose a crear un escenario abierto y crítico que sea apropiado para la puesta en marcha del proceso empoderativo (Fetterman y Wandersman 2007)<sup>41</sup>.

✓ La cultura de evidencia, tal y como se ha anticipado anteriormente, es uno de los pilares de las evaluaciones empoderativas. Su desarrollo exige indagar en los procesos de fundamentación de las distintas posiciones que los actores clave mantienen frente a diferentes eventos y que subyacen de su razonamiento lógico. Una cultura basada en evidencias persigue construir procesos lógicos donde la evidencia empírica o documental juega un rol prioritario de forma que las posiciones basadas en creencias, prejuicios u otro tipo de mecanismos no fundamentados sean progresivamente abandonadas.

✓ Los ciclos de reflexión y acción en las evaluaciones empoderativas se sustentan sobre la lógica de que los datos evaluativos sobre un programa determinado ofrecen información de utilidad para la gerencia, la cual, en el ejercicio de la toma de decisiones que es su responsabilidad, genera nuevos escenarios en el programa evaluado, produciendo nuevos datos evaluativos que nuevamente serán incorporados en los procesos de decisión. El concepto enfatiza la naturaleza cíclica del proceso donde los datos evaluativos son continuamente utilizados como parte del proceso de toma de decisiones y, a su vez, se retroalimentan de dicho proceso como consecuencia de las decisiones adoptadas. Los ciclos de reflexión y acción ponen de manifiesto la naturaleza dinámica de la práctica evaluativa empoderativa la cual está expuesta a una constante evolución. En este contexto, se considera que la evaluación empoderativa puede ser exitosa y generar ciclos de reflexión y acción cuando ésta se encuentra institucionalizada en un organismo determinado y forma parte integral de la planificación y gestión del programa en conexión directa con las instancias de decisión dentro del organismo en cuestión; y difícilmente podrá lograr sus objetivos si se constituye como un elemento aislado.

---

<sup>40</sup>La neutralidad del evaluador/a en la evaluación empoderativa es irrelevante. La evaluación empoderativa comparte el punto de vista de Greene (1997, pág. 25) de que *"los/as evaluadores/as están inevitablemente de un lado, y no de otro. Los bandos elegidos se expresan a través de las cuestiones que se plantean y, por tanto, del criterio utilizado para hacer valoraciones sobre la calidad del programa"*.

<sup>41</sup>El rol del evaluador/a en las evaluaciones empoderativas ha sido objeto de discusión y debate teórico (Smith 2007). Fetterman y Wandersman (2007) se vieron obligados a aclarar que los/as evaluadores/as no son encargados de empoderar a otros, sino de crear el entorno adecuado para que otros se empoderen. Según aseguran, nadie tiene la capacidad de empoderar a otros, el ejercicio de empoderamiento nace de uno mismo.

✓ ‘Comunidad de aprendices’: La evaluación empoderativa facilita la existencia y emergencia de comunidades de aprendices, entendidas éstas como grupos que participan activamente en la evaluación y que están compuestos por individuos que aprenden unos de los otros siendo ellos mismos quienes revisan el trabajo de sus pares. A través de este proceso, la comunidad de aprendices se refuerza al incrementarse la confianza en sus propias capacidades. Asimismo, dado que las personas integradas en estas comunidades de aprendices comparten valores y objetivos comunes, se facilita la cooperación y el apoyo mutuo en la rendición de cuentas.

✓ ‘Actitudes reflexivas: La evaluación empoderativa cultiva actitudes reflexivas al promover activamente la idea de que, a través del análisis conjunto y basado en evidencias, se pueden poner en marcha medidas de mejora aplicables no solo en el programa evaluado sino también en otros ámbitos de la vida laboral y personal.

La metodología de la evaluación empoderativa se subdivide en cuatro etapas (Fetterman 2001):

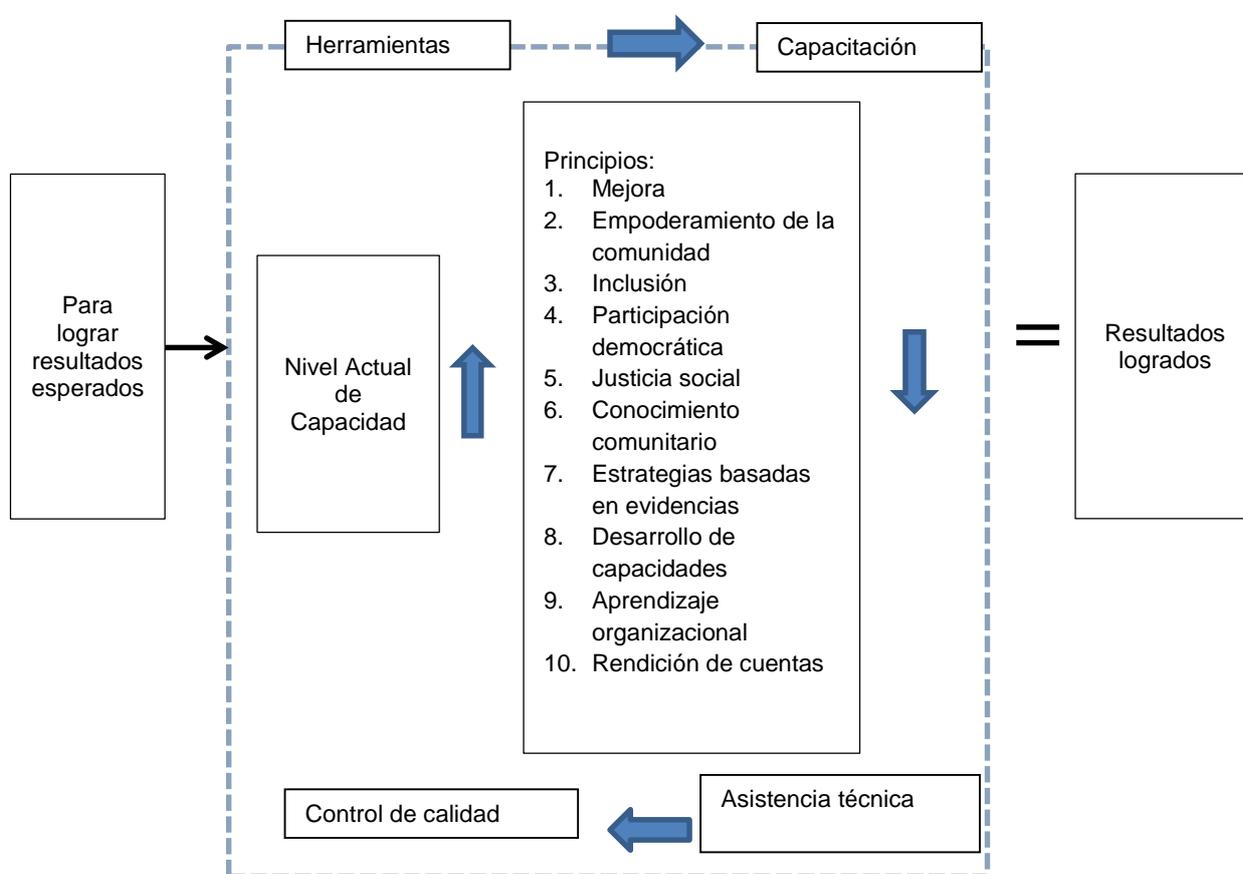
1. Establecimiento de la misión: El/la evaluador/a ayuda al grupo a llegar a un consenso sobre cuál es su misión, su rol, su propósito. Para ello, guía a los y las participantes con la intención de que éstos sugieran ideas sobre dichos elementos, proponiendo así un debate abierto donde se discuten las diferentes propuestas hasta alcanzar un consenso bajo la guía del evaluador/a.

2. Priorización (*Taking stock*): Esta etapa se subdivide en la fase de establecimiento de prioridades y la fase de priorización. El/La evaluador/a guía a los y las participantes en el desarrollo de una lista que incluya las actividades más necesarias e importantes que les permitan cumplir los objetivos programáticos u organizacionales propuestos, los cuales emanan de la fase anterior. Una lista de 20 o 25 actividades puede ser suficiente. Después cada uno de los y las participantes, deberán hacer uso de 5 marcadores con la instrucción de que sean colocados en aquellas actividades que se consideran más importantes. Se pueden poner uno o más marcadores en cada actividad. Las 10 actividades que reúnan un mayor número de marcadores superan esta etapa y pasan a la fase siguiente. En el paso siguiente, para cada una de las actividades seleccionadas, los y las participantes deben ponderar en una escala del 1 al 10 su opinión sobre la situación actual dentro del organismo en cuestión de las actividades seleccionadas. Estas opiniones deben ser sustentadas con datos y evidencias. Los resultados se recogen en un documento que es utilizado en la fase siguiente, ‘planeando el futuro’. En esta etapa, la labor del evaluador/a es actuar de facilitador, instructor y guía.

3. La tercera fase, planeando el futuro, supone la planificación de estrategias, objetivos e indicadores que puedan ser monitoreados posteriormente. Los objetivos tienen que estar relacionados con el análisis del estado de las actividades de la fase anterior.

4. Una vez que se han formulado los objetivos, estrategias e indicadores, comienza la cuarta etapa de monitoreo, la cual implica la utilización de diversas herramientas entre las que se incluyen encuestas, establecimiento de grupos focales, entrevistas y diseños de investigación cuasi-experimentales. Antes de que las herramientas metodológicas mencionadas se pongan en funcionamiento, la propuesta de estrategia de monitoreo se pone a disposición de los y las participantes, quienes realizan aportaciones y aprueban la estrategia final. Entre 3 y 6 meses después de puesto en marcha el sistema de monitoreo, se realiza, de nuevo, una reunión siguiendo la metodología de la fase 2 (*taking stock*). En esta reunión, los y las participantes comparan los resultados que obtuvieron (línea de base) con sus nuevas ponderaciones y ,en base a este análisis, se actualiza el documento formulado originalmente. Este procedimiento se repite sucesivamente en el tiempo, de forma que el ciclo de reuniones y comparaciones no tiene una limitación temporal y se considera abierto hasta que las actividades lleguen a las ponderaciones deseadas o, en su defecto, sean sustituidas por otras nuevas. En muchas ocasiones, se utilizan métodos audiovisuales que reflejen la evolución en las ponderaciones otorgadas, preservando así la memoria organizacional y la validez de las decisiones tomadas (Trochim 2006).

**Figura 2.3. Modelo teórico de la evaluación empoderativa**



Fuente: Fetterman y Wandersman 2007, pág. 193

En opinión de Fetterman, la evaluación empoderativa es un ejercicio de carácter reductivo, ya que se evoluciona desde un amplio espectro de elementos hacia uno más reducido a partir de un enfoque de priorización y pensamiento estratégico. Este proceso ayuda a los y las participantes a ser selectivos y prepararse para la identificación de estrategias y soluciones aplicables a la mejora organizacional. Fundamentalmente, es un ejercicio que responde, a su vez, a un criterio de eficiencia puesto que el proceso delimita el ámbito de actividades a cubrir de forma eficiente y analiza la manera más estratégica de buscar soluciones considerando las fortalezas y debilidades de la organización. Por último, delinea un plan de acción dinámico y sujeto a evolución.

Por todo ello, se puede concluir que en el enfoque de evaluación empoderativa, el desarrollo de la capacidad evaluativa es un principio central (Fetterman et al 1996). Tal y como la propia palabra indica, empodera a las personas que participan en el proceso dejando como legado el fortalecimiento de su capacidad evaluativa.

El análisis de la evaluación empoderativa a través del modelo de Cousins y Whitmore arroja valores máximos de las tres dimensiones de dicho modelo.

Así, el control y la toma de decisiones (primera dimensión) se encuentra completamente en manos de los actores clave; la diversidad de la participación (segunda dimensión) es máxima y puede incluir a grupos de intereses muy diferentes incluyendo donantes y beneficiarios; y el grado de participación (tercera dimensión) es total y extensivo a todas las fases de la evaluación.

#### **2.4.4 La Evaluación Colaborativa**

El término de evaluación colaborativa es el que más debate teórico ha suscitado en el campo de los enfoques participativos de evaluación (Fetterman et al. 2014). Aunque Cousins et al. 2013 utilizan indistintamente el término evaluación colaborativa e investigación colaborativa (*collaborative inquiry*) en general, otros autores (O' Sullivan 2004, Rodríguez-Campos 2005) consideran a la evaluación colaborativa como un enfoque particular y diferenciado dentro de las evaluaciones participativas.

En palabras de O' Sullivan (2007), la evaluación colaborativa es aquella que involucra activamente a los actores clave en la evaluación, pero siempre en la medida en que los actores puedan o quieran colaborar. Similar a los otros enfoques participativos, la evaluación colaborativa también busca 'engancharse' a los actores clave en la evaluación de cara a facilitar el uso de los resultados. Sin embargo, en contraposición a otros enfoques participativos, la evaluación colaborativa se distingue de las demás prácticas evaluativas al otorgar un papel crucial al grado de capacidad y de voluntad que desplieguen los actores clave de un organismo, determinados en base a diferentes escalas de colaboración (O' Sullivan 2004; Rodríguez-Campos 2005).

Los niveles de colaboración deben explorarse y negociarse con los actores clave en la etapa previa a la evaluación. La negociación es así fundamental para el análisis de la capacidad de los actores clave para colaborar en la evaluación y para establecer un plan de recogida de información que permita responder a las preguntas evaluativas. A su vez, esta etapa sirve para fortalecer los conocimientos sobre evaluación de los potenciales actores participantes.

La principal característica de la evaluación colaborativa es, por tanto, su componente negociador para determinar los niveles de colaboración de los actores clave. En este tipo de evaluaciones, el control sobre las decisiones inherentes a la práctica evaluativa está en manos del evaluador/a. En contraste, en la evaluación empoderativa, el control de la toma de decisiones en la evaluación se encuentra en manos de los y las participantes de la misma y tanto en la evaluación basada en la utilización como en la evaluación participativa (práctica o transformativa) se produce un control compartido en la toma de decisiones entre evaluadores/as y actores clave.

Rodríguez-Campos (2005) propone un 'modelo de evaluación colaborativa' donde se destaca la importancia de establecer un equipo de evaluación sólido así como una serie de medidas clave que permitan prevenir potenciales dificultades relativas a este tipo de evaluaciones. El modelo incluye 6 aspectos:

- ✓ Identificación de la situación.
- ✓ Clarificación de las expectativas.
- ✓ Establecimiento de un compromiso compartido.
- ✓ Garantía de comunicación abierta.
- ✓ Promoción de mejores prácticas.
- ✓ Seguimiento de guías específicas.

En la práctica, la evaluación colaborativa tiende a organizarse alrededor de 4 fases, tal y como muestran las distintas experiencias empíricas llevadas a cabo en base a este modelo (Jay et al. 2006, Costelo et al. 2005, O' Sullivan et al. 2005):

1. Revisión del estado del programa: entre los principales aspectos a considerar en esta primera etapa se incluye la profundización en el contexto cultural y político del programa así como su comparación con otros contextos similares. Asimismo, es importante realizar un análisis de evaluaciones previas y la revisión de ejercicios de monitoreo afines. Finalmente, en esta etapa, se incluye un análisis de la cultura evaluativa de la organización, incluyendo los conocimientos del personal sobre evaluación y el nivel de aceptación y receptividad de los actores clave a participar en la evaluación en cuestión.
  
2. Desarrollo del plan de evaluación: esta fase generalmente comienza con una capacitación a los y las participantes sobre los enfoques y métodos del proceso evaluativo. Generalmente, dichas capacitaciones son abiertas y participativas; y en ellas se promueve la identificación de las preguntas evaluativas. Durante este proceso, por tanto, el equipo de evaluación puede involucrar a los actores participantes en la elaboración de un modelo lógico<sup>42</sup> del programa evaluado, intentando que los borradores sean elaborados por ellos mismos (con el apoyo y la facilitación del equipo evaluador). Las experiencias empíricas han demostrado que, en un tercio de los casos donde se aplican evaluaciones colaborativas, los y las participantes son capaces de completar esta actividad sin problemas mientras que en otro tercio de los casos se requiere de varios borradores. En el tercio restante es necesario proveer un apoyo intensivo por parte del equipo evaluador para completar los modelos lógicos.
  
3. Implementación de la evaluación: la fase anterior sirve para identificar las capacidades y el interés de los actores participantes en la evaluación. En la fase de implementación, en cambio, se produce la

---

<sup>42</sup> Modelo lógico entendido como las etapas inherentes a un proceso de cambio, que se representa generalmente a través de un esquema gráfico donde se vinculan todos los elementos inherentes en un programa (actividades, productos, resultados, objetivos).

recogida y el análisis de datos de acuerdo con los métodos de evaluación escogidos en la fase anterior (grupos focales, entrevistas, etc.) y en función del grado de colaboración que se haya acordado con los y las diferentes participantes. Se promueve, a su vez, que los datos recopilados y analizados sean revisados por diferentes grupos de actores para facilitar la transparencia, el control y una amplia implicación en el proceso evaluativo.

4. Diseminación de hallazgos y resultados: esta fase cobra especial relevancia en la evaluación colaborativa y se ha planeado con antelación, durante la elaboración del plan de evaluación, con lo que los y las participantes ya conocen cuáles son sus responsabilidades, responsabilidades que fueron además previamente acordadas. Generalmente, se utilizan diferentes instrumentos para la diseminación como, por ejemplo, eventos, foros y seminarios. Las presentaciones y los informes elaborados son expuestos en dichos eventos y constituyen la base para desarrollar informes posteriores sobre mejores prácticas y/o lecciones aprendidas.

**Figura 2.4. Ciclo de evaluación colaborativa anual**



Fuente: O' Sullivan 2007, pág. 293

Se puede afirmar que este modelo de práctica evaluativa promueve la participación y fortalece las capacidades evaluativas del personal de las organizaciones a través de la colaboración, contribuyendo, por tanto, a

incrementar la capacidad evaluativa organizacional. Igualmente, autores como O' Sullivan y D' Agostino (2002) también han aludido a la reducción de costes como una ventaja destacada de la evaluación colaborativa ya que el número de personal externo necesario para el desarrollo de la evaluación es más limitado. Pero, sobre todo, la evaluación colaborativa encuentra sus más claras ventajas en su habilidad para lograr el empoderamiento de los y las participantes sobre el programa evaluado y sobre la evaluación en sí misma como vehículo para su mejora.

El enfoque colaborativo es probablemente el más difícil de analizar siguiendo el modelo de Cousins y Whitmore propuesto al inicio de este apartado. En cuanto a la dimensión primera de toma de decisiones, se puede afirmar que el evaluador es el que tiene el control de la toma de decisiones, aunque su labor es la de fomentar una fuerte implicación y un acercamiento entre evaluadores/as y actores clave. La diversidad de los actores clave (segunda dimensión) es algo indeterminado, al ser el propio proceso de negociación con estos potenciales participantes el que va a determinar su diversidad sin existir en esta dimensión ningún patrón predeterminado. En relación al grado de intensidad de la participación (tercera dimensión) ocurre algo similar, ya que el grado de colaboración se negocia en la planificación de la evaluación y éste no tiene por qué ser uniforme, pudiendo producirse que la participación se desarrolle en base a diferentes grados y escalas.

Tras esta revisión de los enfoques participativos, se puede afirmar que todos estos modelos han sido reconocidos ampliamente en la literatura evaluadora por su capacidad para incrementar la relevancia y el uso tanto de los procesos evaluativos como de los hallazgos particulares sobre el programa evaluado (Fetterman 1994, Fitzpatrick 1998, Mueller y Fitzpatrick 1998, Ryan y Johnson 2000).

A pesar de presentar diferencias claras entre ellos, algunos/as autores/as se inclinan por el desarrollo de un enfoque participativo sólido y común que establezca de una manera holística unos principios de investigación colaborativa donde todos los enfoques tengan cabida. Esta línea de pensamiento es partidaria de no ahondar tanto en las diferencias existentes entre los diversos modelos participativos (Cousins et al. 2013). Otras perspectivas defienden, sin embargo, la necesidad de seguir profundizando en las diferencias y clarificando los modelos en base a definiciones conceptuales más precisas (Fetterman et al. 2014).

Independientemente de que en este debate nos posicionemos más por la armonización o por la diferenciación entre enfoques participativos, una importante lección aprendida, hasta la fecha, es que las organizaciones, los programas y los proyectos necesitan de recursos, conocimiento y desarrollo organizacional para ser capaces de gestionar sus evaluaciones de manera efectiva y, por tanto, necesitan desarrollar capacidad evaluativa (Patton 2007).

La capacidad evaluativa adquiere nuevamente un rol prioritario, tal y como se ha mencionado anteriormente, al estar estrechamente vinculada con la capacidad de las organizaciones de realizar evaluaciones de calidad y de

conseguir que su utilización sea una práctica rutinaria y sostenible capaz de enriquecer el proceso de toma de decisiones en una organización determinada.

## **2.5. Los modelos de capacidad evaluativa**

Las múltiples definiciones existentes sobre capacidad evaluativa en un contexto organizacional coinciden todas ellas en afirmar que se trata de un proceso dinámico, intencionado, multidisciplinar a través del cual se pretende que la evaluación se convierta en un instrumento de calidad y útil tanto en la toma de decisiones como en el aprendizaje organizacional.

En este apartado se presentan los modelos conceptuales que tratan de analizar exhaustivamente la capacidad evaluativa, desvelando, en particular, sus causas así como los efectos que generan. Éste es sin duda un tema de gran interés y debate actual, que responde a la creciente necesidad de identificar y profundizar en los factores que influyen en el grado de capacidad evaluativa en las organizaciones así como en los resultados y la efectividad de aquellos esfuerzos que, en última instancia, aclaman estar destinados al fortalecimiento de la propia capacidad evaluativa (Baizerman et al. 2002; Carman 2007; Carman y Fredericks 2010; Compton et al. 2002; Cousins et al. 2008; Huffman et al. 2008; Labin et al. 2012; Milstein et al. 2002; Preskill y Boyle 2008; Stevenson et al. 2002; Stockdill et al. 2002).

A continuación se presentan los cuatro modelos más representativos que aparecen en la literatura. Modelos, todos ellos, que tratan de dar respuesta a las siguientes preguntas: ¿cuáles son los factores que facilitan el desempeño de una práctica evaluativa sostenible? ¿Cuáles son las estrategias más adecuadas para obtener dichos factores? ¿Qué aspectos de la cultura organizacional son más relevantes para lograr una transferencia de conocimiento entre los miembros de dicha organización de forma eficaz?

El primer modelo presentado es el Modelo Multidisciplinario de Preskill y Boyle, desarrollado en 2008. Después se profundizará sobre el Enfoque de Inmersión Colaborativa definido por Huffman, Thomas y Laurenz, también en 2008. Ambos modelos han sido aplicados en estudios de caso<sup>43</sup>. El modelo Multidisciplinario ha servido como referente para el desarrollo de otros modelos posteriores como el Marco Interactivo Cultural y Contexto Dependiente de Suarez-Balcázar et al. (2010) y el Modelo Integrado de Capacidad Evaluativa, desarrollado por Labin et al., en 2012.

### **2.5.1 El Modelo Multidisciplinario**

Las autoras Preskill y Boyle proporcionan una serie de elementos que contribuyen a que la capacidad evaluativa pueda cumplir su propósito de hacer de la evaluación una práctica sostenible y que, a su vez, coinciden y desarrollan el concepto de capacidad evaluativa tal y como ha sido definido por Stockdill, Baizerman y Compton (2002).

---

<sup>43</sup>Veáse a tal efecto: A Catalyst for change approach to ECA (García-Iriarte et al. 2011).

Para Preskill y Boyle (2008, pág. 444), el desarrollo de capacidad evaluativa implica la puesta en práctica de una serie de estrategias de enseñanza que ayuden a individuos, grupos u organizaciones a aprender sobre qué elementos constituyen una práctica evaluativa útil, efectiva y profesional. El objetivo último de la capacidad evaluativa es la sostenibilidad de la práctica evaluativa, donde los y las participantes en las actividades destinadas a realizar evaluaciones se preguntan continuamente acerca de cómo recoger, analizar e interpretar datos, así como sobre la forma de usar los hallazgos de las evaluaciones en la toma de decisiones y en la puesta en marcha de medidas concretas. Para que la práctica de la evaluación sea algo sostenible, los actores participantes en la misma deben disponer de un abierto y decidido apoyo gerencial, pero también de incentivos, recursos, y oportunidades que les permitan transferir el aprendizaje que adquieren de la práctica evaluativa a su trabajo diario. La sostenibilidad implica, igualmente, el desarrollo de sistemas, políticas y prácticas que ayuden a integrar el trabajo evaluativo en los procesos de establecimiento de objetivos estratégicos.

Entre los elementos que componen la concepción de capacidad evaluativa de Preskill y Boyle se puede identificar una cierta influencia de los enfoques constructivistas del aprendizaje colaborativo, influencia que se ve reflejada en los siguientes aspectos:

- ✓ La intencionalidad del aprendizaje: Las estrategias de enseñanza y aprendizaje deben facilitar y proveer las herramientas técnicas necesarias para facilitar que las personas miembros de una organización puedan adquirir conocimientos.
  
- ✓ La influencia del contexto de sociabilidad en el conocimiento. Al igual que en los modelos constructivistas, el concepto de capacidad evaluativa asume que solo en el marco de un contexto social determinado se puede lograr un aprendizaje significativo.
  
- ✓ La colaboración. El aprendizaje organizacional posee en sí mismo un carácter colaborativo. Éste alienta así al surgimiento de ideas y prácticas novedosas a partir de la estimulación continua de los integrantes de una organización o colectivo para que intercambien de forma abierta planteamientos y posiciones diferentes (Eggen y Kauchak 1999).

La premisa fundamental del modelo de Preskill y Boyle es que la organización ha de estar comprometida con la internalización de los procesos evaluativos, incluyendo los métodos de evaluación, sus políticas y sus estructuras, elementos todos ellos que de por sí tienen un carácter dinámico. De la misma manera, la organización debe fomentar una cultura donde se valora positivamente la confianza, la asunción de riesgos, la apertura, la curiosidad y la experimentación.

En este modelo, el desarrollo de la capacidad evaluativa es concebido como un proyecto en sí mismo, con sus diferentes fases y objetivos que se centran en el incremento de conocimiento evaluativo dentro de una organización así como en la generación de una actitud más positiva ante la evaluación.

En base a la literatura existente, Preskill y Boyle han identificado un total de 36 objetivos potenciales que pueden servir de guía para los esfuerzos destinados al desarrollo de capacidad evaluativa (ver Tabla 2.2).

Así, Preskill y Boyle enfatizan en la necesidad de elegir los objetivos más adecuados para lograr una mayor capacidad evaluativa teniendo en cuenta el contexto organizacional. Posteriormente, la estrategia que la organización en cuestión ponga en marcha se adecuará a los objetivos seleccionados de forma que se facilite el alcance de los mismos de manera efectiva. Los objetivos potenciales también son la base para determinar el grado de capacidad evaluativa existente en la organización en un momento determinado (línea de base). De hecho, las estrategias seleccionadas para lograr los objetivos propuestos han de tener en cuenta la capacidad ya existente.

La precisión con la que se describen los objetivos de capacidad evaluativa es un elemento relevante del estudio de Preskill y Boyle, ya que estas autoras otorgan una gran importancia a la voluntad y a las motivaciones de las personas miembros de una organización determinada, así como a sus expectativas y asunciones sobre los logros potenciales de la capacidad evaluativa. En otras palabras, los esfuerzos para promover la capacidad evaluativa (enmarcados en los objetivos propuestos) en definitiva son intencionales y de cara a ser efectivos han de ser precisos y realistas.

Por ello, la elección de las estrategias a seguir es un aspecto fundamental en este modelo y el conocimiento del contexto en el que se van a poner en marcha dichas estrategias es muy relevante para hacer bien la elección, al igual que fue relevante, también, para la selección de los objetivos en el paso previo.

Una elección adecuada de la estrategia a seguir incrementa sustancialmente la habilidad de la organización de alcanzar el objetivo de desarrollo de capacidad evaluativa previamente propuesto; y el logro de éste, no solo generaría un incremento de la capacidad evaluativa sino que sería una prueba en sí mismo de un acto de aprendizaje organizacional. Como se observaba al inicio de este capítulo, el aprendizaje organizacional es uno de los propósitos principales de la capacidad evaluativa y la consecución de los objetivos propuestos, a partir de la implementación de estrategias acertadas, representa en última instancia signos de aprendizaje organizacional.

**Tabla 2.2. Objetivos en el desarrollo de capacidad evaluativa**

<b>Conocimiento (cognitivo)</b>
<p>Los actores participantes en las actividades de capacidad evaluativa entienden:</p> <ul style="list-style-type: none"> <li>• Que la evaluación es un conjunto de actividades planificadas, sistemáticas y con un propósito determinado.</li> <li>• Los conceptos y términos evaluativos.</li> <li>• La relación entre investigación y evaluación.</li> <li>• Cómo los procesos y hallazgos evaluativos pueden contribuir a la toma de decisiones.</li> <li>• Las fortalezas y debilidades de los diferentes enfoques en evaluación.</li> <li>• Las fortalezas y debilidades de los diferentes métodos de obtención de información.</li> <li>• Cómo realizar análisis estadísticos a partir de datos cuantitativos.</li> <li>• Cómo realizar análisis temáticos a partir de datos cualitativos.</li> <li>• Cómo los procesos y hallazgos evaluativos pueden estar afectados por consideraciones políticas.</li> <li>• La importancia de usar métodos y enfoques evaluativos de manera responsable y en consideración con el contexto y la cultura organizacional.</li> <li>• Los aspectos éticos en la práctica evaluativa.</li> <li>• Que diferentes actores claves pueden tener opiniones, perspectivas y experiencias diferentes con respecto a una evaluación.</li> <li>• La relación existente entre los objetivos de un programa, con sus actividades y resultados.</li> <li>• Los requisitos importantes a tener en cuenta en la contratación de un evaluador/a.</li> </ul>
<b>Habilidades (comportamiento)</b>
<p>Los actores participantes en las actividades de capacidad evaluativa son capaces de:</p> <ul style="list-style-type: none"> <li>• Desarrollar un modelo lógico.</li> <li>• Desarrollar preguntas evaluativas.</li> <li>• Escribir un plan de evaluación.</li> <li>• Diseñar métodos de obtención de información.</li> <li>• Elegir métodos de obtención de información apropiados y relevantes.</li> <li>• Recoger información creíble y comprobable.</li> <li>• Analizar datos cuantitativos.</li> <li>• Analizar información cualitativa.</li> <li>• Interpretar resultados que lleven a hallazgos.</li> <li>• Desarrollar el presupuesto para una evaluación.</li> <li>• Comunicar y describir el proceso evaluativo y sus hallazgos a través de estrategias variadas.</li> <li>• Aplicar los estándares de evaluación.</li> <li>• Enseñar a otros sobre evaluación.</li> <li>• Desarrollar una plan estratégico evaluativo.</li> <li>• Gestionar un proceso evaluativo.</li> </ul>
<b>Creencias (afectivo)</b>
<p>Los actores participantes en las actividades de capacidad evaluativa creen que la evaluación:</p> <ul style="list-style-type: none"> <li>• Proporciona información útil.</li> <li>• Puede resultar una experiencia positiva.</li> <li>• Debe formar parte del proceso del diseño del programa.</li> <li>• Contribuye al éxito del programa.</li> <li>• Añade valor a la organización.</li> <li>• Es una parte importante de su trabajo.</li> <li>• Compensa tanto en tiempo como en costos empleados.</li> </ul>

Fuente: Elaboración propia, adaptada de Preskill 2007, págs. 226-227

Preskill y Boyle proponen un conjunto de diez estrategias que pueden ser utilizadas para el alcance de los objetivos previamente establecidos, y que se agrupan, a su vez, en objetivos vinculados al conocimiento evaluativo, a las habilidades técnicas de evaluación y a actitudes sobre la práctica evaluativa. Estas estrategias son:

- ✓ La implicación en el proceso de evaluación: en el modelo multidisciplinario de creación de capacidad evaluativa, la participación emerge como vehículo principal para el aprendizaje organizacional, sin hacer valoraciones acerca de cuál de los enfoques participativos revisados en el apartado anterior es el más adecuado.
- ✓ Capacitación: la forma, duración y metodología empleada para capacitar a las personas miembros de una organización puede ser diversa y no está predeterminada en el modelo multidisciplinario. Pueden incluirse tanto cursos formales sobre evaluación como experiencias de aprendizaje de corto plazo donde se usan una variedad de actividades para responder a las necesidades de los participantes (Brooks-Harris y Stock-Ward 1999).
- ✓ Apoyo técnico: se refiere a la importancia de tener la posibilidad de recibir el apoyo de un/a evaluador/a, bien sea externo o interno. El modelo adopta el enfoque de Yin y White (1984) donde el apoyo desde un punto de vista técnico permite una mejor adopción e implementación de prácticas o procedimientos pedagógicos. Esto implica tener un mayor acceso a redes y comunidades profesionales en evaluación y, por tanto, genera oportunidades para solucionar problemas específicos (Fruchter, Cahill y Wahl 1998).
- ✓ Materiales escritos: incluye el desarrollo de material didáctico sobre evaluación, como manuales, boletines e incluso otro tipo de documentación más específicos, que refleja planes e informes de evaluación concretos, con un enfoque orientado al aprendizaje.
- ✓ Comunidades de práctica: en general, se refiere a aquellos grupos de personas que comparten información, experiencias y metodologías en un área común de interés. Su composición y establecimiento puede ser formal o informal, dentro de la empresa, o incluso entre organizaciones.
- ✓ Investigación apreciativa: es considerada parte del enfoque colaborativo de aprendizaje y se centra en el análisis de los elementos que han sido exitosos dentro de la organización, tratando así de desvelar cuáles son las fortalezas en la misma. Whitney y Trosten-Bloom (2003, pág. 1) puntualizan que *“la investigación apreciativa es el estudio y la exploración de qué es lo que hace que los sistemas organizacionales funcionen de una manera óptima. Ésta se basa en la*

*premisa de que el diálogo sobre las fortalezas, valores y esperanzas tienen un carácter transformativo. La investigación apreciativa de acuerdo al autor sugiere así que el cambio organizativo parte de un proceso de investigación, afirmación y apreciación”.*

✓ Tecnología: El uso y adaptación de fuentes conectadas a internet (*on line*) como páginas web, blogs, etc. pueden ser útiles en el desarrollo de la capacidad evaluativa. En este campo, el abanico de opciones es amplio y puede incluir la creación de bases de datos sobre evaluaciones, repositorios de material evaluativo o incluso portales con listados de páginas web especializadas.

✓ Pasantías (*Intershops*): se considera positivo para el desarrollo de capacidades evaluativas fomentar la colaboración entre estudiantes y equipos de evaluación en el marco de diferentes tareas de investigación, preparación de material educativo o incluso la impartición de cursos sobre evaluación.

✓ Tutorías (*Mentoring o coaching*): supone el establecimiento de una relación sólida y duradera del personal de una organización con profesionales de la evaluación a través de la cual se dispone de apoyo individualizado especialmente en el ámbito del desarrollo profesional (Davis 2001).

✓ Reuniones: La planificación de encuentros donde se discutan actividades relacionadas con la evaluación para aprender y compartir experiencias es considerada como un factor importante en el modelo de capacidad evaluativa multidisciplinar.

Como se ha mencionado anteriormente, la identificación de los objetivos y de las estrategias correspondientes es un ejercicio secuencial que busca la implementación óptima de un proceso de desarrollo de capacidad evaluativa (considerado por las autoras como la demanda de capacidad evaluativa). Adicionalmente, el modelo de Preskill y Boyle, también propone una serie de recursos que debe ser garantizados por el organismo en cuestión, y que consideran muy necesarios para que las estrategias puedan ser efectivas de cara a alcanzar los objetivos, y hacer así de la evaluación una práctica sostenible (oferta de la organización en materia de capacidad evaluativa):

✓ Políticas y procedimientos evaluativos: facilitan la institucionalización de la evaluación en la organización, mediante la creación o el establecimiento de unidades de evaluación o la designación de personal responsable de facilitar el desarrollo de capacidades evaluativas en la organización. Se incluyen aquí también aquellos esfuerzos dirigidos a la introducción de procesos en las prácticas y rutinas organizacionales que permitan que la evaluación

permea en el trabajo diario del personal de la organización (Compton et al. 2002; King 2002; Volkov y King 2007).

✓ Guías y manuales sobre evaluación: clarifican expectativas y valores que serán la base para construir los métodos y enfoques evaluativos de la organización. Permiten un acercamiento a los estándares evaluativos adaptados al lenguaje y al contexto de la organización en cuestión, lo que facilita un mayor entendimiento y conocimiento evaluativo y, por ende, una mayor sostenibilidad de su práctica (Milstein et al. 2002).

✓ Existencia de recursos dedicados a la evaluación: la realización de evaluaciones de calidad requiere de recursos financieros, humanos y materiales (Volkov y King 2007). De la misma manera, los esfuerzos para el desarrollo de la capacidad evaluativa también requieren de recursos. Es decir, hay que dedicar personal (externo o interno) a la puesta en marcha de las estrategias y actividades y proporcionar una tecnología adecuada a dichas estrategias (software, bases de datos, materiales audiovisuales, etc.).

✓ Uso de hallazgos procedentes de evaluación: de cara a ser sostenible, la evaluación debe utilizar todas las oportunidades a su alcance para promocionar su uso, no solo en el momento de tomar decisiones, sino también en la práctica diaria a través de diversas actividades de comunicación.

✓ Compromiso y receptividad sobre la práctica evaluativa: los miembros de una organización deben creer en el valor añadido de la evaluación y en lo que ésta puede aportar al cumplimiento de las diferentes metas organizacionales establecidas. Este hecho se refleja en la opinión de las personas sobre la evaluación y de su interés en aplicar herramientas evaluativas en sus rutinas diarias (Huffman et al. 2006).

✓ Sistema integrado de gestión del conocimiento: la práctica sostenible de la evaluación depende, en gran medida, de la capacidad de la organización de capturar, analizar y transformar la información en conocimiento. Esto implica la integración de la evaluación en el sistema de conocimiento de la organización en cuestión, sincronizando la actividad evaluativa con el ciclo programático y generando así complementariedades de flujos de información a diferentes niveles.

✓ Plan estratégico de evaluación: La existencia de un plan estratégico de evaluación es relevante para clarificar los motivos y objetivos de la evaluación, así como otros aspectos fundamentales de la misma (qué evaluar, porqué, cómo, cuándo y con qué destinatarios). Un plan estratégico ayuda a la organización a tener una actitud proactiva y a

considerar la evaluación dentro de un ciclo organizacional, lo que contribuye a la sostenibilidad de la función de evaluación.

✓ Aprendizaje continuo sobre evaluación: Ante posibles rotaciones de personal que pongan en peligro la pérdida del conocimiento evaluativo adquirido a través del desarrollo de capacidad evaluativa, las organizaciones deben proporcionar oportunidades de aprendizaje continuo, con actividades variadas y complementarias que faciliten su replicabilidad y sostenibilidad.

Por último, el modelo aborda los elementos catalizadores en una organización que, en última instancia, se convierten en determinantes de su capacidad de aprendizaje. El grado de existencia de estos elementos influye sustancialmente en la conexión entre la demanda de la organización (estrategias y objetivos para fomentar el desarrollo capacidad evaluativa) y la oferta que ésta realiza en esta esfera (recursos para una evaluación sostenible):

✓ Cultura, que refleja las tradiciones, valores y elementos sustanciales que ayudan a comprender los patrones de conducta de las personas integrantes de una organización. La cultura de una organización que promueve el desarrollo de la capacidad evaluativa es proclive a la experimentación y al cambio (Preskill y Torres 1999; Rowe y Jacobs 1998; Torres y Preskill 2001; Toulemonde 1999).

✓ Liderazgo: es necesario un claro y determinante apoyo gerencial para la implementación y sostenibilidad de la capacidad evaluativa (Cousins et al. 2004; Goh y Richards 1997; King 2002; Milstein et al. 2002; Owen y Lambert 1995).

✓ Sistemas y estructuras: las estructuras existentes en las organizaciones son las que permiten que el personal interactúe, colabore y se comunique en base a un sistema. El éxito de los esfuerzos destinados al desarrollo de la capacidad evaluativa dependerá, en gran medida, de la flexibilidad de los sistemas y estructuras establecidos ya que parte del personal deberá compatibilizar sus actividades diarias con otras actividades relacionadas con la participación en tareas evaluativas (Torres y Preskill 2001).

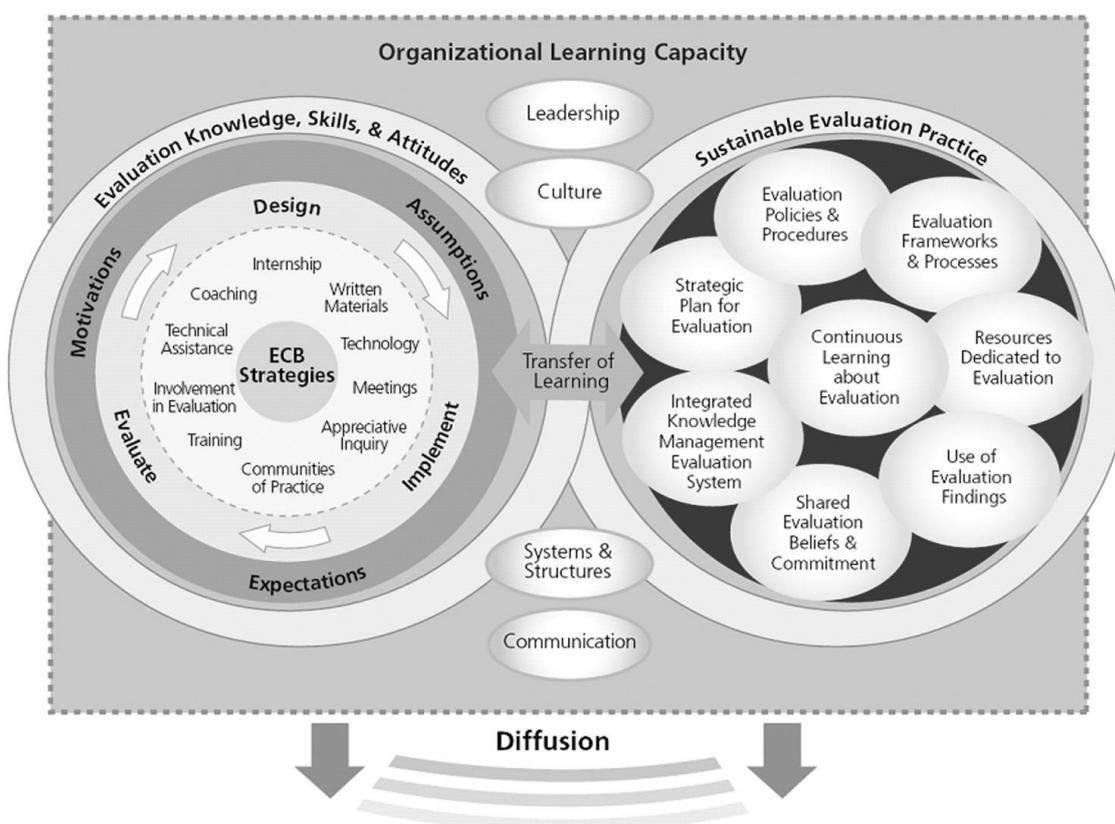
✓ Entorno: en ocasiones la necesidad de rendir cuentas a agentes externos actúa como factor motivante del desarrollo de la capacidad evaluativa en las organizaciones (Stockdill et al. 2002; Sutherland 2004; Toulemonde 1999).

El grado de existencia de estos cuatro factores internos determina finalmente el nivel de capacidad de aprendizaje que una organización posee y, por tanto,, la habilidad de dicha organización para que su demanda de capacidad evaluativa

(objetivos y estrategias) case con su oferta de capacidad evaluativa (recursos internos). En el caso de que la demanda sea satisfecha por la oferta, la información fluiría con mayor facilidad en un escenario de transferencia de conocimiento y aprendizaje (*knowledge transfer and learning*). El aprendizaje es, por tanto, esencial en este modelo ya que en definitiva lo que se persigue es promoverlo a través de la práctica evaluativa.

La figura 2.5 (a continuación), describe el Modelo Multidisciplinario de Preskill y Boyle.

**Figura 2.5. Modelo Multidisciplinario**



Fuente: Preskill y Boyle 2008, pág. 3

### 2.5.2 El Enfoque de Inmersión Colaborativa

Al igual que en el Modelo Multidisciplinario explicado anteriormente, Huffman, Thomas y Laurenz (Huffman et al. 2008) parten de la definición de capacidad evaluativa propuesta por Stockdill et al (2002) y proponen un nuevo modelo al que denominan Enfoque de Inmersión Colaborativa. Este modelo persigue resolver una importante limitación detectada en el modelo multidisciplinario en relación a la fragilidad y dudosa idoneidad de algunas de las estrategias que propone. Por lo tanto, la principal aportación del enfoque de inmersión colaborativa va dirigida a superar esta limitación, adoptando elementos del enfoque colaborativo de evaluación participativa (ver apartado 2.2.4).

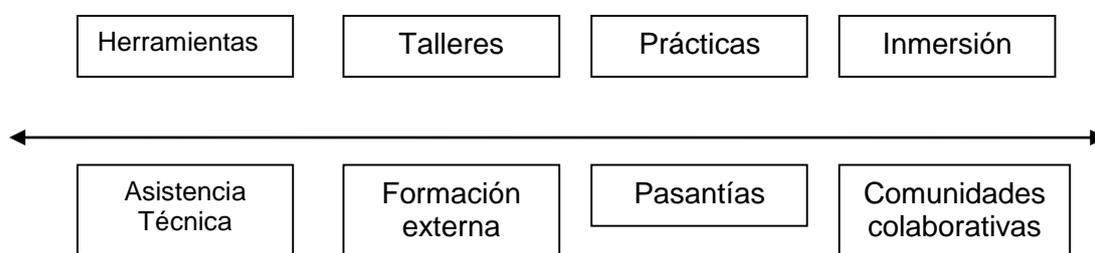
En consecuencia, el enfoque de inmersión colaborativa tiene importantes similitudes con el modelo multidisciplinario de Preskill y Boyle anterior y mantiene estrategias como las capacitaciones, la asistencia técnica, el *coaching/mentoring*, etc. Sin embargo, el enfoque de inmersión colaborativa enfatiza la necesidad de favorecer y fomentar la inmersión, la participación y la colaboración de personas y grupos de interés en las tareas evaluativas como estrategia fundamental para desarrollar capacidad evaluativa.

Sus reminiscencias con los enfoques constructivistas de aprendizaje son claras y reconocidas por sus autores. De hecho, desde el enfoque de inmersión colaborativa se enfatiza la necesidad de que, para poder desarrollar capacidad evaluativa, las personas de una organización tengan la oportunidad de construir su propio entendimiento sobre los conceptos evaluativos en base a los conceptos teóricos y herramientas adquiridas (racionalismo) así como a la experiencia (empirismo).

A su vez, las personas y grupos de la organización deben tener la oportunidad de desarrollar y aplicar los conceptos evaluativos que mejor encajen en su contexto organizacional. Se ha de disponer del tiempo y de la libertad para considerar y discutir las implicaciones de sus acciones, enmarcadas en un entorno que sea proclive al cambio. Además, en este proceso colaborativo las personas participantes de la organización deben poder contar con el apoyo y la guía de expertos/as que faciliten su aprendizaje y su comprensión de los conceptos evaluativos.

En consecuencia, como ya se ha adelantado, la principal característica de este modelo está en la estrategia propuesta para el desarrollo de capacidad evaluativa, una estrategia de naturaleza claramente colaborativa. Este enfoque, en línea con las perspectivas de aprendizaje constructivistas, puede desarrollar de manera más efectiva el conocimiento y las habilidades evaluativas en las personas y, simultáneamente, construir capacidad evaluativa en la organización. La figura 2.6 describe los tipos de estrategias propuestas bajo el enfoque de inmersión colaborativa.

**Figura 2.6. Estrategias en el enfoque de inmersión colaborativa**



Fuente: Elaboración propia, adaptado de Huffman et al. 2008, pág. 361

Como se puede observar en la figura 2.6, en el nivel de aprendizaje más elemental, la capacidad evaluativa puede fortalecerse a través de la adquisición individual y ad hoc de conocimiento, proveniente de eventuales consultas a expertos/as o del análisis de material específico sobre temas de naturaleza

evaluativa. En el nivel siguiente de aprendizaje, la capacidad evaluativa puede facilitarse a través de la participación en actividades más estructurales como talleres o cursos formativos donde las personas retornan a la organización a la que pertenecen con los nuevos conocimientos adquiridos. En un nivel más avanzado de aprendizaje, las personas no solo participan en cursos, sino que trabajan directamente en evaluaciones de otras organizaciones. Esta estrategia proporciona la posibilidad de enfrentarse a experiencias reales en el ámbito evaluativo que el aprendizaje a través de talleres o cursos no ofrece. El cuarto nivel de aprendizaje es el vinculado directamente a la propuesta realizada por el enfoque de inmersión colaborativa. En este nivel, los individuos participan activamente en múltiples evaluaciones de su propia organización, guiados por personas expertas que pueden ser externas o internas a la organización. Los expertos/as ayudan a clarificar los conceptos y proporcionan apoyo sustantivo en los diferentes análisis que se realizan. La idea final de este enfoque es facilitar que las personas de una organización puedan encontrarse inmersas en el desarrollo e implementación de una serie de evaluaciones donde, a través de su implicación y cooperación, aprendan y compartan experiencias y conocimientos.

Este enfoque es muy similar a los enfoques de evaluación participativos, en especial al colaborativo y al de evaluación empoderativa, pero con la particularidad de que estos enfoques de evaluación participativos no conciben la participación como parte de una concepción sistemática de la práctica evaluativa. La principal motivación de los modelos participativos es así la utilización eficiente de los hallazgos evaluativos a través de la participación, mientras que en el enfoque de inmersión colaborativa, el objetivo es que las personas se constituyan en promotores de la capacidad evaluativa de la organización donde trabajan a través de su participación sistemática en evaluaciones. De esta forma, la inmersión en la práctica evaluativa se convierte en su principal fuente aprendizaje.

### **2.5.3 El Marco Interactivo Cultural**

El Modelo Interactivo Cultural<sup>44</sup> fue desarrollado por Suarez-Balcázar et al. (2010) con el fin de profundizar en aquellos factores que son considerados cruciales para el desarrollo de la capacidad evaluativa. Para ello, en su propuesta se realiza una revisión sistemática de la teoría y de la práctica existente en materia de capacidad evaluativa y, en particular, su análisis se centra en la premisa teórica de que existen determinados factores, individuales y organizacionales, que son ciertamente más proclives a desarrollar capacidad evaluativa y, por tanto, a facilitar el logro de aquellos objetivos vinculados con el desarrollo de esta capacidad. El elemento distintivo de este marco teórico es, sin duda, la idea de que los factores individuales que incrementan la capacidad evaluativa interactúan con factores organizacionales y ambos, a su vez, están inmersos en la cultura y el contexto de la organización (Flashpohler et al. 2008).

---

<sup>44</sup> En inglés, *Cultural and Contextual Framework for Capacity Building*

El modelo de Suarez-Balcázar et al. se asienta en los trabajos conceptuales sobre capacidad evaluativa desarrollados por Baizerman et al. (2002) y Stockdill et al. (2002), así como en los diversos análisis realizados sobre el desarrollo de capacidades a nivel comunitario (Chaskin 2001; Flashpohler et al. 2008; Lord y Hutchison 2003; Nye y Glickman 2000; Ristau 2001). La principal aportación de este modelo es la identificación de una serie de factores facilitadores o predictores de capacidad evaluativa que, además, el modelo presenta, categorizados de acuerdo a su aplicación individual u organizacional. Esta clasificación de factores facilitadores en individuales u organizativos también está presente en los trabajos de Carman y Fredericks 2008; Cousins et al. 2004; García-Iriarte et al. 2011 y Nielsen et al. 2011.

De acuerdo a Suarez-Balcázar et al., los **factores individuales** desencadenantes de capacidad evaluativa son tres:

- ✓ La predisposición y motivación para implicarse activamente en tareas evaluativas son dos aspectos clave para absorber cualquier contenido didáctico en esta área y, por tanto, para utilizar la evaluación de forma sostenible (Stockdill et al. 2002). No es sólo una cuestión de motivación, sino también de compromiso y de conciencia de que la evaluación puede ser útil si ésta se integra en la práctica diaria. Tal y como subrayan Fawcett et al. (2003) y Edwards et al. (2000), dentro de una organización pueden coexistir distintos niveles de predisposición individual y, a través de prácticas como la propia capacitación, se pueden lograr aumentos en los niveles de predisposición para participar en prácticas evaluativas.
  
- ✓ Las competencias (conocimiento y habilidades): este segundo factor individual se refiere al conocimiento sobre evaluación disponible así como a las habilidades evaluativas adquiridas desde un punto de vista técnico: por ejemplo, la capacidad para seleccionar las preguntas clave de una evaluación o para desarrollar un modelo lógico (Kaplan y Garrett 2005), o incluso para analizar e interpretar datos de diversa índole (Fawcett et al. 2003). Existen diferentes maneras de transferir las competencias evaluativas entre los individuos dentro de una organización. Así, pueden utilizarse ejercicios de capacitación, prácticas de intercambio temporal de personal entre departamentos o la disponibilidad de grupos de apoyo que faciliten asistencia técnica puntual. Las actividades para fomentar la transferencia de competencias y habilidades forman parte de una estrategia intencional y continua (Fawcett et al. 1994, 1996).
  
- ✓ El conocimiento del contexto y la competencia cultural individual: el personal que va a participar en el proceso de desarrollo de capacidad evaluativa debe conocer el contexto de la organización, sus estructuras, sus relaciones de poder y sus prácticas, puesto que este conocimiento, sin duda, va a jugar un rol importante en el proceso de desarrollo de capacidad evaluativa (Stockdill et al. 2002).

Los **factores organizacionales** que contribuyen al desarrollo de capacidad evaluativa son aquellos que proporcionan la estructura necesaria para optimizar los esfuerzos dirigidos a dicho objetivo. El modelo de Suarez-Balcázar et al. identifica cuatro factores organizacionales clave:

✓ Liderazgo: este es un factor estrechamente relacionado con la efectividad de cualquier estrategia de capacidad evaluativa en las organizaciones (Kettner et al. 1999; Sobeck y Agius 2007). Un liderazgo dirigido a promover el desarrollo de la capacidad evaluativa debe conseguir que la evaluación sea una función importante e independiente dentro de la organización, motivando a su personal en el desarrollo de habilidades evaluativas (Flashpohler et al. 2008; Greenhalgh et al. 2004). El papel señalado que el liderazgo desempeña en la sostenibilidad de las habilidades evaluativas del personal de la organización, así como en la motivación del personal a la hora de utilizar los hallazgos de las evaluaciones para la mejora de los programas y la rendición de cuentas, es un elemento clave identificado en diferentes estudios sobre la materia (Goodman et al. 1998; Alaimo 2008; Milstein et al. 2002). La existencia de líderes comprometidos con la evaluación es esencial para lograr que la evaluación sea parte rutinaria de las actividades de una organización (Flashpohler et al. 2008). Preskill y Torres, a su vez, también incluyen en su modelo el liderazgo como uno de los elementos intrínsecos que apoyan y facilitan el aprendizaje organizacional. Taut (2007), por su parte, subraya que el férreo apoyo del liderazgo dentro una organización es absolutamente necesario para lograr que la evaluación sea valorada como herramienta de aprendizaje en la misma. Kotter (1996) y Milstein et al. (2002) también apoyan estas tesis que sitúan al liderazgo como pilar del desarrollo de capacidad evaluativa.

✓ Clima de aprendizaje: actualmente existe un amplio consenso en el reconocimiento que se le otorga al aprendizaje como vehículo fundamental para que una organización pueda alcanzar sus objetivos. Inmersas en la continua búsqueda de mejoras tanto en el servicio que se ofrece como en la calidad del mismo, las organizaciones son cada vez más conscientes de la importancia que tiene instaurar un clima de aprendizaje dentro de sus estructuras y tareas. En este sentido, Herman y Gioia (1998, pág 12) subrayan que *“cuando las personas aprendemos se produce un crecimiento en múltiples ámbitos, incluyendo progreso laboral, satisfacción personal y autoestima”*. De acuerdo con los autores mencionados, las organizaciones que invierten en el aprendizaje de su personal se posicionan con mayores perspectivas de éxito.

✓ Recursos en apoyo a la evaluación: como cualquier otra actividad con capacidad de producir cambios organizacionales e influir en la toma de decisiones, tanto la práctica evaluativa como el desarrollo de la capacidad evaluativa, requieren de recursos materiales, financieros y humanos (Arnold 2006; Volkov y King 2007). De acuerdo a Stockdill et

al. (2002) el desarrollo de capacidad evaluativa es un proceso intencional y no accidental, donde se requiere de una amplia gama de recursos humanos y financieros, que incluyen personal, espacio físico dentro de las organizaciones, tiempo invertido, equipos informáticos, sistemas de documentación etc.

✓ Cultura organizacional: de nuevo en base a la definición de Baizerman et al. (2002), la capacidad evaluativa depende del contexto y debe adaptarse a la forma en la que la organización opera. Las organizaciones se caracterizan por tener una determinada cultura que afecta no solo al comportamiento de sus integrantes y a las pautas de relacionamiento entre ellos, sino también a los valores, normas y tradiciones que respetan y promulgan (Suarez-Balcázar 2002). En este modelo, el factor cultural es ampliamente considerado para el desarrollo de capacidad evaluativa.

Este modelo ha sido validado empíricamente por Taylor-Ritzler et al. (2013), quienes han analizado los factores individuales y organizacionales que comprenden la capacidad evaluativa y que inciden en una mejor utilización de los hallazgos evaluativos. En su estudio, las autoras desarrollaron un cuestionario donde los factores individuales y organizacionales fueron desglosados en un total de 68 preguntas (indicadores), en una escala de valoración que abarcaba desde 'totalmente de acuerdo', a 'totalmente en desacuerdo'. Cerca de 850 organizaciones no gubernamentales del área de Chicago fueron invitadas a participar, y se obtuvieron 169 respuestas. Los resultados de la encuesta corroboran la relación existente entre factores individuales y organizacionales y sus efectos directos en los resultados de capacidad evaluativa. En particular, los resultados sirvieron para reafirmar el rol crucial que juegan los factores organizacionales sobre la efectividad de los factores individuales en el uso de la evaluación.

Asimismo, el estudio señalaba que, incluso cuando las personas integrantes de una organización poseen el conocimiento adecuado y la motivación necesaria para hacer de la evaluación una práctica rutinaria, este cambio tiene menos posibilidades de producirse cuando la organización no convergen las premisas previamente analizadas de liderazgo (*leadership*), clima de aprendizaje y recursos destinados a la evaluación (Taylor-Ritzler et al. 2013).

#### 2.5.4 El Modelo Integrador de Capacidad Evaluativa

Este modelo ha sido desarrollado por Labin et al. en 2012, a partir de un análisis del marco teórico y de los casos empíricos publicados hasta la fecha, con el fin de identificar aspectos comunes en el desarrollo de la capacidad evaluativa y facilitar la selección de estrategias adecuadas.

En este trabajo, se parte de una revisión documental que refleja la forma en la que distintas organizaciones están implementando estrategias de desarrollo de capacidad evaluativa. En particular, además de las estrategias, la investigación ofrece una amplia gama de información relativa al desarrollo de capacidad evaluativa, como los resultados obtenidos y el contexto en el que se desenvuelven las acciones para lograr una mayor capacidad evaluativa.

El modelo integrador presenta un importante vínculo con los enfoques participativos, colaborativos y empoderativos de evaluación. De acuerdo a Labin et al. (2012), la evaluación empoderativa y otros enfoques participativos y colaborativos de evaluación han realizado importantes contribuciones a la consolidación del desarrollo de capacidad evaluativa ya que han demostrado desde la propia práctica evaluativa que objetivos similares a los propuestos para el desarrollo de la capacidad evaluativa pueden ser obtenidos a partir de promover una amplia participación.

El modelo propone una definición de capacidad evaluativa que se asienta conceptualmente en las definiciones comúnmente ya aceptadas así como en los modelos teóricos propuestos por Preskill y Boyle (2008). Así, se define capacidad evaluativa como *“todo aquel proceso intencional que persigue incrementar la motivación, el conocimiento y las herramientas que permitan fortalecer la habilidad de un grupo o de una organización para realizar o utilizar evaluaciones”* (Labin et al. 2012, pág. 308).

La principal aportación de este modelo integrador es la aplicación del denominado ‘marco lógico’ en su versión básica de Necesidades-Actividades-Resultados aplicado a la capacidad evaluativa (Kellogg Foundation 2001; United Way 1996, 2008). El marco lógico implica una dirección causal de izquierda a derecha a partir de la premisa lógica de que las necesidades indican las actividades y éstas, a su vez, los resultados obtenidos (Figura 2.7).

**Figura 2.7. Modelo Integrador de Capacidad Evaluativa**

Necesidades	Actividades	Resultados
<ul style="list-style-type: none"> <li>• Motivaciones                             <ul style="list-style-type: none"> <li>○ Audiencia (Interna o externa)</li> <li>○ Asunciones</li> <li>○ Expectativas</li> </ul> </li> <li>• Objetivos</li> <li>• Contexto                             <ul style="list-style-type: none"> <li>○ Análisis de necesidades</li> <li>○ Adaptación</li> </ul> </li> <li>• Recursos y fortalezas                             <ul style="list-style-type: none"> <li>○ Individuales</li> <li>○ Organizacionales                                     <ul style="list-style-type: none"> <li>▪ Recursos</li> <li>▪ Experiencia</li> <li>▪ Prácticas, liderazgo, cultura</li> </ul> </li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Estrategias                             <ul style="list-style-type: none"> <li>○ Teoría</li> <li>○ Modo</li> <li>○ Nivel individual u organizacional</li> </ul> </li> <li>• Implementación                             <ul style="list-style-type: none"> <li>○ Meta                                     <ul style="list-style-type: none"> <li>▪ De población</li> <li>▪ Organización</li> <li>▪ Dominio</li> </ul> </li> <li>○ Tiempo, frecuencia</li> <li>○ Correcciones a medio termino</li> <li>○ Barreras</li> </ul> </li> <li>• Evaluación de las actividades de capacidad evaluativa                             <ul style="list-style-type: none"> <li>○ Enfoque</li> <li>○ Diseño</li> <li>○ Mediciones</li> <li>○ Tipo de información</li> <li>○ Marco temporal</li> <li>○ Interna o externa</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>▪ Individuales                             <ul style="list-style-type: none"> <li>○ Actitudes</li> <li>○ Conocimiento</li> <li>○ Habilidades</li> </ul> </li> <li>▪ A nivel de la organización                             <ul style="list-style-type: none"> <li>○ Procesos, políticas y practicas</li> <li>○ Liderazgo</li> <li>○ Cultura</li> <li>○ Enfoques</li> <li>○ Recursos</li> </ul> </li> <li>▪ Resultados</li> <li>▪ Lecciones aprendidas</li> </ul>

Fuente: Elaboración propia, adaptada de Labin et al. 2012, pág. 209

La primera columna del modelo se refiere al propósito o motivación que a guiado a la organización a emprender actividades de desarrollo de la capacidad evaluativa, las necesidades (Preskill y Boyle 2008). En este sentido, la motivación puede tener un origen interno a la organización, externo o proceder de una combinación de ambos. El análisis de la motivación de la capacidad evaluativa puede facilitar la identificación de objetivos y el desarrollo de estrategias adecuadas para la organización. Así, conocer la actitud y el compromiso hacia la evaluación, los recursos destinados a la capacidad evaluativa, la experiencia previa con la práctica evaluativa y otros aspectos organizacionales, como el apoyo por parte de la gerencia hacia la evaluación y, en general, la cultura organizacional, pueden ayudar a identificar e implementar las estrategias más adecuadas (Owen 2003; Preskill y Boyle 2008; Suarez-Balcázar et al. 2010). Estos factores mencionados inciden positivamente en el aprendizaje organizacional y por tanto en el grado de efectividad y sostenibilidad de la capacidad evaluativa (Preskill y Boyle 2008). En el Modelo Integrador de Capacidad Evaluativa estas características preexistentes se mencionan como fortalezas y recursos de la organización de cara al desarrollo de la capacidad evaluativa.

La segunda columna del Modelo Integrador de Capacidad Evaluativa señala las estrategias y actividades específicas destinadas al fortalecimiento evaluativo. Al igual que Preskill y Boyle (2008), en este modelo se destaca la importancia de introducir los enfoques participativos y las teorías de aprendizaje y de cambio organizacional a la hora de la identificación de actividades y estrategias para el desarrollo de capacidad evaluativa. Las estrategias así pueden incluir actividades diversas como talleres de capacitación, la provisión de asistencia técnica en el área evaluativa o la promoción del involucramiento y participación directa de personal de la organización en actividades evaluativas (Duffy y Wandersman 2007).

Labin et al. presentan una particularidad adicional en su modelo y es la atención que se presta en este modelo a la necesidad de monitorear y aplicar medidas correctoras en el desarrollo e implementación de las estrategias seleccionadas para lograr un mayor fortalecimiento evaluativo (Wandersman et al. 2000). Los esfuerzos destinados a fortalecer la capacidad evaluativa de las organizaciones deben incluir, por tanto, mecanismos para evaluar los resultados de dichos esfuerzos y reconducirlos cuando esto sea necesario. Para ello, es importante desarrollar indicadores, medirlos y analizar el progreso efectivo en la consecución de los objetivos marcados (Preskill y Boyle 2008).

El modelo integrador señala, también, una serie de factores que son cruciales para poder lograr los resultados esperados de incremento de capacidad evaluativa. Estos factores pueden ser individuales o pueden referirse a un nivel organizacional.

En primer lugar, si nos referimos a los resultados a nivel individual, se destaca como factor primordial la necesidad de que se apoye la existencia de un contexto organizacional favorable y proclive para el desarrollo de capacidad evaluativa. La razón reside en que normalmente los resultados esperados a nivel individual incluyen la mejora en la actitud ante la evaluación y un mayor conocimiento y habilidades evaluativas (Duffy y Wandersman 2007; Owen 2003; Preskill y Boyle 2008; Suarez-Balcázar et al. 2010). Para lograr los resultados mencionados, Preskill y Boyle subrayan que el aprendizaje individual está enormemente influenciado por el contexto organizacional y Taylor-Ritzler et al. (2013) encontraron, en su estudio de caso, que el apoyo organizacional es, a su vez, necesario para el aprendizaje individual y para generar cambios. Un contexto organizacional favorable emerge así como un factor fundamental a nivel individual.

En segundo lugar, a nivel organizacional, el Modelo Integrador de Capacidad Evaluativa subraya cinco características importantes para lograr el éxito de las estrategias que persiguen el fortalecimiento de la capacidad evaluativa:

- ✓ Existencia de procesos, políticas y prácticas para la realización y el uso de la evaluación (GAO 2003; Suarez-Balcázar et al. 2010).
- ✓ Existencia de un sólido liderazgo en esta área siendo este un aspecto importante para lograr un cambio organizacional (Kotter 1996).

- ✓ Cultura organizacional favorable, entendida como la acumulación de valores, actitudes, objetivos y prácticas que puedan apoyar el cambio organizacional (Boyle 1999; Preskill y Torres 1999; Suarez-Balcázar et al. 2010).
- ✓ Convertir la evaluación en una práctica rutinaria y considerarla como un elemento esencial en la sostenibilidad de la capacidad evaluativa (Boyle 1999; Preskill y Boyle 2008; Sanders 2003; Stockdill et al. 2002)
- ✓ Destinar recursos a la práctica evaluativa, siendo este un factor importante para el éxito y sostenibilidad de la capacidad evaluativa (Boyle et al 1999; Gibbs et al 2002; Preskill y Boyle 2008).

Como hemos visto en esta revisión documental en relación a la capacidad evaluativa, la literatura existente se enfoca en dar respuesta a cuatro preguntas clave: ¿qué es capacidad evaluativa?, ¿cuál es su impacto?, ¿cómo podemos conocer el grado de capacidad evaluativa que existe en una organización?, y ¿cómo podemos hacer que la capacidad evaluativa de una organización aumente ?

A pesar de que se han producido múltiples esfuerzos de investigación en torno a la capacidad evaluativa, casi toda la literatura analizada coincide en señalar la necesidad de seguir profundizando aún mas en el estudio y el análisis empírico en esta materia y, en particular, en la búsqueda de más evidencia empírica que valide los modelos conceptuales existentes y que demuestre que los esfuerzos destinados al fortalecimiento de la capacidad evaluativa son justificables y tienen un impacto positivo en la buena marcha de los programas y, en definitiva, de las organizaciones (Preskill y Boyle 2008; Suarez-Balcázar et al. 2010).

De esta manera, hasta la fecha, la literatura teórica y empírica se basa frecuentemente en la premisa de que la capacidad evaluativa facilita el uso de las evaluaciones, su calidad y, por consiguiente, la mejora del rendimiento organizacional en general. Sin embargo, con la excepción de algunos estudios recientes (como el desarrollado por Taylor-Ritzler et al. 2013), existe escasa evidencia empírica que demuestre que las prácticas efectuadas para el desarrollo de la capacidad evaluativa en las organizaciones hayan incidido 'de facto' en una mejora en el uso de la evaluación, en su calidad, y en la institucionalización de la práctica evaluativa en sí misma. Queda, por tanto, un amplio camino por recorrer en el área del desarrollo de capacidad evaluativa, materia de estudio que necesita de más información sistematizada que la fortalezca e incluso que la justifique.

Este trabajo de investigación tiene el propósito de acercar aún más la experiencia práctica al desarrollo teórico y se asienta sobre el convencimiento de que, en la esfera de la capacidad evaluativa, es tan necesario que los

esfuerzos de investigación se nutran de experiencias prácticas como que las organizaciones se impregnen de los nuevos conocimientos desarrollados para poder beneficiarse de ellos.

### **3. La evaluación en el Sistema de las Naciones Unidas**

La revisión en detalle del origen y evolución de la disciplina evaluativa realizada en los dos capítulos anteriores refleja la importancia que ha tenido para el desarrollo teórico, la acción de la política pública y la necesidad de valorar los méritos de ésta. La política pública (vinculada en gran parte a los estados y administraciones de ámbito doméstico) se ha constituido así como punto central del desarrollo de la evaluación.

No obstante, esta tesis se centra en el Sistema de las Naciones Unidas (NU) fundamentalmente porque es un área de acción internacional de gran trascendencia global que ha adoptado de forma decisiva el marco teórico alrededor de la práctica evaluativa al entender la evaluación como un vehículo muy importante para el aprendizaje y, en última instancia, para la mejora del desempeño de las instituciones que componen el sistema. De esta forma, existe un amplio consenso interno y externo sobre la idea de que, a partir de la evaluación, las instituciones de las NU pueden servir mejor a los objetivos de mantenimiento de paz, desarrollo sostenible y reducción de la pobreza, entre otros, que los países miembros han trazado.

En este sentido, esta tesis se asienta sobre el convencimiento de que, desde el ámbito de las NU, también se puede contribuir a construir un marco teórico sólido y fundamentado empíricamente en cuanto a sus proposiciones en materia de evaluación. El Sistema de las NU representa hoy a un conjunto de organismos internacionales que movilizan una gran cantidad de recursos y que en la práctica se han unido a los esfuerzos de desarrollo de capacidad evaluativa para mejorar el propio desempeño de las instituciones que lo componen. Así, desde esta tesis se comparte la idea de que las enseñanzas y lecciones aprendidas, que se derivan de éstos esfuerzos, que tienen lugar en el ámbito internacional pueden contribuir a consolidar aún más el marco teórico de esta disciplina.

El análisis que se realiza en este capítulo permite, por tanto, contextualizar la hipótesis principal de esta tesis en el Sistema de las NU. Recordar que esta hipótesis principal establece que un elemento clave dentro de las instituciones cómo es la cultura de aprendizaje, si viene acompañado de un sólido marco normativo, entonces se constituye en un factor fundamental para el logro de la capacidad evaluativa, que permite, a su vez, la mejora de la calidad de las evaluaciones y su influencia en la toma de decisiones. Esta combinación de factores posee una importancia no reconocida en los modelos teóricos que actualmente analizan la capacidad evaluativa y el objetivo de este trabajo es reforzar dichos modelos teóricos de forma que los hallazgos que esta tesis ofrezca puedan aplicarse no sólo en el ámbito de las NU sino también en otras áreas de acción circunscritas al ámbito estatal o regional.

En este contexto, el apartado primero de este capítulo examina la práctica de la evaluación en el Sistema de las NU, empezando por el estudio de las diferentes formas posibles de desempeño de la función evaluativa en los organismos de las NU desde un punto de vista funcional. Posteriormente, se estudian en detalle las políticas y procedimientos evaluativos más comunes

dentro del Sistema de las NU, incluyendo el marco referencial de Normas y Estándares de UNEG. Finalmente, en un tercer apartado, se revisan los esfuerzos realizados en materia de desarrollo de capacidad evaluativa, profundizando en la justificación de dichos esfuerzos y en sus principios rectores.

### **3.1. El Sistema de las NU: la función de evaluación y el objeto evaluado**

Como se ha señalado brevemente en el capítulo anterior (apartado 2.2.), el Sistema de las NU emerge después de la Segunda Guerra Mundial con la finalidad de evitar conflictos entre países y de preservar la paz a nivel mundial. Así, en 1945, 50 países firman y ratifican la Carta de las Naciones Unidas, la cual se constituye como el documento referencial del gran número de organismos y agencias especializadas que emergen en su entorno con objetivos afines aunque diferenciados.<sup>45</sup>

De esta manera, a partir de la ratificación de la Carta de las Naciones Unidas y del surgimiento progresivo de agencias, programas y organizaciones inspirados por sus principios y valores, el Sistema de las Naciones Unidas se conforma como una estructura internacional de carácter gubernamental. Ésta se compone por un conjunto de organismos con diferentes mandatos vinculados a objetivos como el ya destacado de preservar la paz a nivel global (el cual es de póstuma importancia en los orígenes del sistema), pero también otros objetivos como promover el desarrollo sostenible de los pueblos y naciones, erradicar los niveles de extrema pobreza, lograr la equidad de género y proveer ayuda humanitaria en situaciones de crisis. Aunque de gran relevancia, éstos son solo algunos ejemplos de las múltiples metas perseguidas por el amplio número de instituciones que conforman el Sistema de las NU.

Más específicamente, el Sistema de las NU se compone actualmente de un total de 15 agencias especializadas, 11 programas y fondos, 6 institutos de investigación y 4 organizaciones relacionadas<sup>46</sup>, todos ellos sujetos el amparo y guía de seis órganos de gobierno donde 193 países miembros<sup>47</sup> se encuentran representados (la Asamblea General, el Consejo de Seguridad, el Consejo Económico y Social, el Consejo de Administración Fiduciaria, la Secretaría y la Corte Internacional de Justicia)<sup>48</sup>.

Esta heterogeneidad de instituciones y mandatos se ve reflejada en la diversidad de mecanismos de gestión y coordinación que se pueden encontrar dentro de las NU, diversidad que también se observa en la práctica evaluativa. De esta forma, podemos distinguir a los organismos de acuerdo a la forma en que han absorbido la función de evaluación dentro de sus estructuras para mejorar el desempeño de sus actividades regulares.

---

<sup>45</sup> Ver [http://www.un.org/en/aboutun/history/sanfrancisco\\_conference.shtml](http://www.un.org/en/aboutun/history/sanfrancisco_conference.shtml)

<sup>46</sup> Ver <http://www.un.org/es/sections/about-un/funds-programmes-specialized-agencies-and-others/index.html>

<sup>47</sup> Ver el sitio <http://www.un.org/en/members/growth.shtml>

<sup>48</sup> [http://www.un.org/es/aboutun/structure/org\\_chart.shtml](http://www.un.org/es/aboutun/structure/org_chart.shtml)

Dentro del Sistema de las NU, además nos encontramos con dos entidades dedicadas exclusivamente a la supervisión y en cuyo mandato es central la propia actividad evaluativa: la Dependencia Común de Inspección (*Joint Inspection Unit, JIU*) y la Oficina de Servicios de Supervisión Interna (*Office of Internal Oversight Services, OIOS*). Ambas disponen de presupuesto propio y línea directa de reporte a los órganos de gobierno del propio Sistema de las NU (la Asamblea General y la Secretaría respectivamente).

En el resto de organismos con mandatos específicos como los mencionados anteriormente, la función de evaluación puede encontrarse más o menos institucionalizada. Así, encontramos casos donde la función de evaluación goza de gran independencia y autonomía, desempeñándose ésta desde departamentos que superan las líneas de gerencia internas y reportan directamente a los órganos de gobierno de su propio organismo<sup>49</sup>. Sin embargo, en la mayoría de los casos, la función de evaluación se desempeña – bien en solitario o junto a otro tipo de funciones como la de investigación o la de auditoría - desde unidades que siendo independientes de los departamentos temáticos, no lo son de la gerencia interna y reportan directamente al máximo responsable del organismo en cuestión<sup>50</sup>. También existen casos donde la unidad de evaluación forma parte de departamentos que reportan a niveles gerenciales intermedios<sup>51</sup> o incluso casos donde la función evaluativa es muy incipiente y dispone de una capacidad de evaluación limitada<sup>52</sup>.

Igualmente, antes de abordar en detalle los procedimientos evaluativos dentro del Sistema de las NU, los cuales vienen en buena parte determinados por las diferentes maneras de absorción de la función de evaluación, es importante describir brevemente el objeto evaluado y las representaciones que éste puede tener dentro del sistema, al constituirse este punto como otro elemento determinante de la forma y fondo de la práctica de evaluación.

A diferencia del ámbito gubernamental estatal, en la esfera de la cooperación internacional al desarrollo y en particular de las NU, los programas y proyectos se constituyen como una unidad de intervención frecuente (objeto evaluado), aunque no son la única. Así, desde los diferentes organismos se realizan esfuerzos de muy diversa índole sujetos a diversas rutas evaluativas:

- ✓ Misiones de paz y seguridad: implican el despliegue de un amplio rango de recursos incluyendo en ocasiones fuerzas militares de las Naciones Unidas con el objetivo más frecuente de preservar la paz en países o regiones inestables, denominadas operaciones de mantenimiento de paz (*peacekeeping operations*). En este ámbito concurren una amplia gama de operaciones que se extienden desde acciones de prevención de conflictos (normalmente organizadas alrededor de amplios esfuerzos

---

<sup>49</sup>Por ejemplo, el Fondo para el Medioambiente Mundial (FMAM) y agencias especializadas como las pertenecientes al Grupo del Banco Mundial.

<sup>50</sup>La mayoría de organismos se encuentran en este grupo, como por ejemplo el PNUD, UNICEF, FAO, UNO Mujeres etc.

<sup>51</sup>Algunos ejemplos: UNICRI, UNCTAD

<sup>52</sup>Entre ellas se encuentran DS, OAJ, OAD

diplomáticos) hasta operaciones de re-establecimiento de paz (cuyo objetivo es retornar a una situación pacífica donde el conflicto ha estallado), bien de forma voluntaria o a través de medidas coercitivas ante una autorización explícita del Consejo de Seguridad).<sup>53</sup> OIOS juega un papel clave en la evaluación de este tipo de acciones, emitiendo un informe de evaluación de las mismas con una periodicidad anual.<sup>54</sup>

- ✓ Acciones de asesoramiento dentro de un país o región (*advocacy*). Este tipo de acciones son frecuentemente el resultado del amplio despliegue de agencias especializadas y programas de desarrollo de las NU en los países en vías de desarrollo, quienes a través de sus representantes y personal especializado realizan una labor de asesoramiento constante a diferentes niveles de gobierno, labor que, en última instancia, persigue influir en la acción de política pública del país o región en cuestión para lograr objetivos de desarrollo sostenible y de reducción de la pobreza.
- ✓ Acciones de fortalecimiento del marco jurídico normativo de un país o mejora de sus capacidades en determinadas áreas jurídicas, legislativas, productivas, educativas, etc.
- ✓ Asistencia financiera y/o técnica en obras de infraestructura básica o productiva.

Las acciones arriba mencionadas, salvo aquellas relacionadas con intervenciones en áreas de conflicto, se organizan frecuentemente en el marco de proyectos y programas<sup>55</sup>, los cuales representan unidades de intervención claramente definidas en cuanto a su duración, presupuesto, actividades y objetivos se refiere. Así, la actividad evaluativa en las NU implica, con frecuencia, diseñar e implementar evaluaciones de programas y proyectos, aunque también se desempeñan evaluaciones que superan el ámbito programático y se centran en áreas temáticas concretas (ej. salud, educación), evaluaciones sectoriales (por ejemplo, centradas en evaluar la acción o estrategia de las NU en determinados sectores productivos) o evaluaciones circunscritas a zonas geográficas determinadas (por ejemplo, evaluaciones de país).

---

<sup>53</sup> Ver <http://www.un.org/en/peacekeeping/operations/peace.shtml>

<sup>54</sup> Ver <https://oios.un.org/page?slug=reports-to-governing-bodies>

<sup>55</sup>En esta tesis los términos de programa y proyecto se utilizan de forma conjunta para reflejar un tipo específico de intervención programática que se caracteriza por ser finita y definida en cuanto a sus plazos, presupuestos, actividades y objetivos. Igualmente se reconoce que ambos términos presentan diferencias. Así, el concepto de programa se vincula frecuentemente a intervenciones de mayor alcance que los proyectos en varias de sus dimensiones, como alcance de sus objetivos, área geográfica e incluso duración temporal. En múltiples ocasiones, un programa incluye el agregado de una serie de proyectos en un ámbito determinado. Igualmente, algunos autores como Alonso y Mosley (1999) relacionan los programas a intervenciones vinculadas al ámbito de la ayuda alimentaria y a la ayuda financiera a través de apoyo a la balanza de pagos (por ejemplo, con condonación de deuda) y ayuda presupuestaria. Lógicamente las evaluaciones de programas en múltiples ocasiones exige acudir al ámbito macroeconómico para, a través de diversas herramientas estadísticas, buscar relaciones causales entre las contribuciones del programa y variables tales como crecimiento, déficit fiscal, inversión, etc.

### **3.2. Formas de evaluación en el Sistema de las NU**

Una vez señalados los dos factores principales que influyen sustancialmente en la forma en que la práctica evaluativa se desempeña en el Sistema de las NU (grado de institucionalización de la función de evaluación y tipo de objeto evaluado), este apartado se centra en describir más específicamente cómo se desempeña dicha función evaluativa en las diferentes instancias dentro del sistema.

#### **3.2.1 Dependencia Común de Inspección (Joint Inspection Unit, JIU)**

La DCI es el único órgano externo independiente de supervisión del Sistema de las NU al que se le ha conferido el mandato de realizar evaluaciones, inspecciones e investigaciones en todo el ámbito de los organismos de las NU.

Esta unidad fue creada de manera experimental por la Asamblea General a través de la resolución 2150 de 1966. Después de dos extensiones aprobadas por la Asamblea General, finalmente en 1976 adquiere el estatus formal como un órgano subsidiario de la misma a través de la resolución 31/192 del 22 de Diciembre de 1976. Así, el JIU responde a la Asamblea General y también a los órganos de gobierno de las agencias especializadas, fondos, programas y organismos afines al Sistema de las NU que han aceptado sus estatutos.

Actualmente, 28 organizaciones han aceptado los estatutos de JIU (JIU 2014), lo que significa que bajo su supervisión se incluyen fondos y programas, agencias especializadas y la Secretaría. Así, la DCI es el órgano de supervisión de mayor alcance dentro del Sistema las NU abarcando la gran mayoría de instituciones que lo componen. Su mandato es también amplio y persigue los siguientes objetivos:

- ✓ Prestar asistencia a los órganos legislativos de las organizaciones participantes en el cumplimiento de sus responsabilidades de gobernanza respecto de la función de supervisión relacionada con la gestión de los recursos humanos, financieros y de otra índole que llevan a cabo sus respectivas secretarías.
- ✓ Contribuir a mejorar la eficiencia y eficacia de las secretarías respectivas en el cumplimiento de los mandatos legislativos y los objetivos de misión establecidos para las organizaciones.
- ✓ Promover una mayor coordinación entre las organizaciones del Sistema de las Naciones Unidas.
- ✓ Determinar las prácticas más idóneas así como proponer marcos de referencia que faciliten el intercambio de información en todo el sistema.

En contraste con lo anterior, la DCI es de reducidas dimensiones en cuanto a personal y recursos. Ésta se compone actualmente de 11 inspectores/as nombrados por la Asamblea General la cual les asigna amplias facultades en

los asuntos relativos a actividades de las NU y, en especial, en lo concerniente a la eficiencia de los servicios y al uso adecuado de los recursos. En relación a su financiamiento, el presupuesto de la JIU se satisface de forma repartida entre las organizaciones participantes de acuerdo al contrato respectivo de cada una de ellas con el organismo. En el 2012, el presupuesto regular ascendió a 6,4 millones de US\$.<sup>56</sup>

A su vez, la DCI persigue proporcionar revisiones independientes en diferentes áreas del Sistema de las NU con el propósito de mejorar la gestión y la coordinación entre organizaciones. En este sentido, los informes que produce, aunque son dirigidos a las organizaciones a las cuales pertenecen las actividades evaluadas, son también compartidos con todas las organizaciones cuando el objeto evaluado es de interés general, ofreciendo así la posibilidad de considerar los resultados de la evaluación a los órganos legislativos de los organismos que componen todo el Sistema de las NU. En el caso de que se trate de documentos confidenciales, éstos son sometidos a los máximos responsables de los organismos para su uso discrecional.

La DCI produce aproximadamente entre 6 y 10 informes anuales así como – aún en menor medida - notas y cartas de gestión confidencial; y sus temas varían desde aspectos generales que afectan al Sistema de las NU en su totalidad, hasta cuestiones específicas relativas a una organización en particular. En su ejercicio de planificación anual, la DCI decide su plan de trabajo e incluye en él, los temas y ejercicios que abordará en el año correspondiente en base a criterios y observaciones propias del equipo de inspección así como a solicitudes provenientes de las organizaciones participantes. En este ejercicio de planificación también se incluyen consultas y sugerencias de otros organismos encargados de la supervisión del sistema, como OIOS o la Junta de Auditores, (*Board of Auditors*), así como de los máximos responsables de los organismos de las NU.

Debido fundamentalmente a la limitada capacidad de gestión y recursos de la DCI, el Sistema de las NU también dispone de otros mecanismos para hacer evaluaciones sobre temas que superan el interés particular de un organismo determinado del sistema. Entre ellos se destacan los siguientes:

- ✓ Las evaluaciones requeridas de manera *ad hoc* por parte tanto de la Asamblea General o de la Secretaría General y que son llevadas a cabo desde la propia Secretaría o por departamentos específicos vinculados a la misma (por ejemplo, paneles de alto nivel establecidos para evaluar la gobernabilidad y la coherencia del Sistema de las NU, la reforma del Sistema de las NU, diversas operaciones de paz, etc.). Estas evaluaciones son asignadas a la agencia más apropiada de acuerdo a la afinidad de su mandato con el objeto evaluado.

---

<sup>56</sup>Para un mayor detalle en relación a las proporciones de cada organismo, ver <https://www.unjiu.org/es/about-us/Paginas/Funding--Budget.aspx>, así como el informe "Report of the Joint Inspection Unit for 2014 and programme of work for 2015" disponible en [https://www.unjiu.org/en/corporate-information/AR%20%20PoW/A\\_69\\_34\\_English.pdf](https://www.unjiu.org/en/corporate-information/AR%20%20PoW/A_69_34_English.pdf)

- ✓ Casos específicos como la revisión de la denominada Política Integral Cuatrienal<sup>57</sup>, la cual evalúa la efectividad y la eficiencia de la asistencia del Sistema de las NU dirigida a fortalecer los esfuerzos de los países en desarrollo por alcanzar sus prioridades de desarrollo y hacer frente a sus necesidades en el contexto de conferencias internacionales como la Declaración del Milenio entre otras. Esta evaluación es realizada por el Departamento de Asuntos Sociales y Económicos de las Naciones Unidas (UNDESA) y es sometida a la Asamblea General y al Consejo Económico y Social para su consideración.

### **3.2.2 La Oficina de Servicios de Supervisión Interna (OIOS)**

OIOS fue establecida por la Asamblea de las NU a través de la resolución 48/218B de 1994 con el propósito de asistir a la Secretaría General de las NU en sus funciones de supervisión y control de sus recursos. Para ello, OIOS se constituyó como una oficina independiente con tres diferentes áreas de control: auditoría Interna, investigaciones y evaluación e inspecciones.

Funcionalmente, OIOS está situada en el organigrama de las NU como una unidad dependiente de la Secretaría General, prestando servicios de supervisión a los organismos y departamentos que dependen del Secretario General. OIOS cubre así a las cuatro secretarías de NU en Nueva York, Ginebra, Viena y Nairobi, a las comisiones regionales, a las operaciones de paz y operaciones humanitarias, y a los Fondos y Programas administrados bajo la autoridad de la Secretaría General (PNUMA, UNO-Habitat, ACNUR).

La labor de OIOS consiste principalmente en asegurar que las actividades implementadas por los organismos mencionados no solo se ajustan a las normas y regulaciones vigentes sino también a los programas aprobados por los propios organismos en los cuales se enmarcan dichas actividades. En este escenario, OIOS valora los méritos de dichos programas y proyectos y recomienda la adopción de medidas orientadas hacia una implementación eficiente y efectiva de los mismos.

Además de trabajar intensamente en el área programática, OIOS también se centra en acciones para prevenir y detectar el fraude, así como el mal uso de recursos, el abuso o la mala gestión. A su vez, OIOS destina importantes esfuerzos a detectar y destacar todas aquellas buenas prácticas y lecciones aprendidas que pueden servir de guía y orientación dentro del sistema. A través de sus informes, el trabajo de OIOS se orienta entorno a la promoción de un sistema de rendición de cuentas que sea transparente y efectivo, y en el fortalecimiento de las capacidades de la Secretaría para identificar y mitigar riesgos que puedan dificultar el logro de objetivos y resultados.

En particular, OIOS realiza las siguientes acciones:

- ✓ Proporciona información y análisis independientes que faciliten la toma de decisiones

---

<sup>57</sup><http://www.un.org/en/ecosoc/about/qcpr.shtml>

- ✓ Realiza análisis independientes de la efectividad en el uso de los recursos
- ✓ Propone medidas para asistir a las diferentes entidades, fondos y programas en cómo responder a riesgos y beneficiarse de oportunidades
- ✓ Promueve una cultura de cambio, orientado a resultados, y gestión efectiva.

En relación a su gobernabilidad interna, OIOS está dividida fundamentalmente en 3 secciones: auditoría interna, investigaciones y evaluación e inspecciones y su máximo responsable tiene rango de Sub-Secretario/a General, y reporta directamente al Secretario/a General quien las pone a consideración de la Asamblea General (Quinto Comité).

OIOS coordina estrechamente sus actividades con la Dependencia Común de Inspección y la Junta de Auditores a través de reuniones regulares.

Los recursos aprobados por la Asamblea General para OIOS en el bienio 2014-2015 ascienden a 121,000 Millones USD<sup>58</sup>. Por secciones, el departamento de evaluaciones e inspecciones tiene asignado un 20% del presupuesto, frente al 38% que va destinado a auditoría interna, y el 27% a investigaciones. El restante 15% se destina a gastos de dirección y apoyo a programas. Existe así un alto nivel de detalle y transparencia en relación a los recursos asignados para cada actividad de OIOS, puesto que, en su mayor parte, están incluidos en el presupuesto ordinario de la Secretaría.

En cuanto a la función de evaluación específicamente, ésta se estableció formalmente el 1 de enero de 2008 en seguimiento a los acuerdos adoptados entre los estados miembros y la Secretaría de las NU en la Cumbre Mundial de 2005<sup>59</sup> sobre la reforma de la gestión de las NU y, en particular, el fortalecimiento de su capacidad de supervisión. Su mandato queda establecido en el Boletín del Secretario General ST/SGB/2000/8<sup>60</sup> en el que se establece que la función de evaluación “*debe contribuir a determinar de manera sistemática y objetiva la relevancia, eficiencia, efectividad e impacto de las actividades de un organismo determinado en relación con sus objetivos, con el propósito de incrementar la efectividad de los programas modificando su contenido, y si es necesario, revisando sus objetivos*” (ST/SGB/2000/8, pág. 12).

La sección de evaluación e inspecciones adquiere, por lo tanto, un papel ciertamente sustantivo dentro de OIOS, e inicia, a partir de su integración formal en 2008, el desarrollo de evaluaciones programáticas así como evaluaciones temáticas e inspecciones generalmente sobre procesos

---

<sup>58</sup> A/68/6

<sup>59</sup> Ver [http://www.un.org/spanish/summit2005/fact\\_sheet.html](http://www.un.org/spanish/summit2005/fact_sheet.html)

<sup>60</sup> Accesible en

<http://daccess-dds-ny.un.org/doc/UNDOC/GEN/N00/408/48/PDF/N0040848.pdf?OpenElement>

específicos y/o unidades dentro de un organismo determinado. A su vez, esta sección dentro de OIOS realiza de manera bianual el informe denominado "*Fortalecer el rol de evaluación y la aplicación de los hallazgos evaluativos en el diseño de programas, asistencia y las directivas políticas*". Por su importancia en la revisión de la capacidad evaluativa en las organizaciones este informe se explica con más detalle en el apartado 3.4.2.

El plan de evaluación de OIOS es preparado en base a un criterio de riesgo. Este criterio asegura transparencia en la selección de áreas u organismos, a la vez que relevancia y puntualidad de las evaluaciones. Es aprobado por el Quinto Comité de la Asamblea General.

### **3.2.3 La evaluación en los organismos de Secretaría General**

Dada la heterogeneidad de las actividades de los diferentes organismos que reportan a la Secretaría General<sup>61</sup>, además de las evaluaciones realizadas por OIOS, dichos organismos también tienen establecidas funciones de evaluación propias, aunque el tamaño y grado de consolidación de dichas funciones de evaluación puede variar de manera considerable de unas a otras. Estos son organismos que forman parte de la Secretaría General:

- ✓ Oficina de Asuntos Jurídicos (OAJ)
- ✓ Departamento de Asuntos Políticos (DAP)
- ✓ Oficina de Asuntos de Desarme de las Naciones Unidas (OAD)
- ✓ Departamento de Operaciones de Mantenimiento de la Paz (DOMP)
- ✓ Departamento de Apoyo a las Actividades sobre el Terreno (DAAT)
- ✓ Oficina de Coordinación de Asuntos Humanitarios (OCAH)
- ✓ Departamento de Asuntos Económicos y Sociales (DAES)
- ✓ Departamento de la Asamblea General y de Gestión de Conferencias (DAGGC)
- ✓ Departamento de Información Pública (DIP)
- ✓ Departamento de Seguridad (DS)
- ✓ Departamento de Gestión (DG)
- ✓ Órganos de justicia interna
- ✓ Otras oficinas: Oficina de Apoyo a la Consolidación de la Paz, Caja Común de Pensiones del Personal de las Naciones Unidas, Secretaría de la Junta de los jefes ejecutivos del sistema de las Naciones Unidas

---

<sup>61</sup> Los Fondos y Programas PNUMA, UNO Habitat y ACNUR, que reportan a la Secretaría, se han incluido en el apartado 3.2.4

para la coordinación, Oficina de Ética, Oficina del Pacto Mundial, Oficina de las Naciones Unidas para las Asociaciones, Fondo de las Naciones Unidas para la Democracia, Personal de la Secretaría de las Naciones Unidas, Servicio de Enlace de las Naciones Unidas con las Organizaciones no Gubernamentales.

- ✓ Asesores, Representantes y Enviados especiales: Oficina del Alto Representante para los Países Menos Adelantados, los Países en Desarrollo Sin Litoral y los Pequeños Estados Insulares en Desarrollo; Oficina del Asesor Especial para África; Oficina del Asesor Especial del Secretario General sobre la Prevención del Genocidio; Oficina del Representante Especial del Secretario General para la Cuestión de los Niños y los Conflictos Armados; Oficina de las Naciones Unidas sobre el Deporte para el Desarrollo y la Paz; Representantes Especiales y Personales y Enviados del Secretario General
- ✓ Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (ACNUDH)
- ✓ Oficina de las Naciones Unidas para la Reducción del Riesgo de Desastres
- ✓ El Organismo de Obras Públicas y Socorro de las Naciones Unidas para los Refugiados de Palestina en el Cercano Oriente (OOPS)
- ✓ La Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD)
- ✓ ONU-Mujeres
- ✓ Oficina de las Naciones Unidas en Ginebra (ONUG)
- ✓ Oficina de las Naciones Unidas en Nairobi (ONUN)
- ✓ Oficina de las Naciones Unidas en Viena (ONUUV)
- ✓ Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC)
- ✓ Comisión Económica y Social para Asia y el Pacífico (CESPAP)
- ✓ Comisión Económica y Social para Asia Occidental (CESPAO)
- ✓ Comisión Económica para África (CEPA)
- ✓ Comisión Económica para Europa (CEE)
- ✓ Comisión Económica para América Latina y el Caribe (CEPAL)

Como se ha mencionado anteriormente, a pesar de su diversidad, la práctica de la evaluación forma parte del quehacer de estos organismos en la mayoría

de los casos, aunque se observan notables diferencias entre ellos. La tabla 3.1 describe la estructura organizacional de la función de evaluación en los organismos de la Secretaría.

**Tabla 3.1. Estructura de la función de evaluación en los organismos de la Secretaría**

Unidad Independiente	Unidad de evaluación integrada en división multifuncional	Unidad no dedicada a evaluación e integrada en división multifuncional	No hay unidad de evaluación, pero existe actividad de evaluación	Actividad mínima de evaluación o prácticamente inexistente
UNODC	DOMP	DAGGC	DAES	DS
ONU Mujeres	CEPA	DIP	DG	OAJ
	CESPAP	CEPAL	DAP	UNUG
	OCAH	CESPAO	CEE	ONUN
	OOPS	ACNUDH	OARPMA	ONUJ
		UNCTAD	OASA	

Fuente: Elaboración propia, adaptada de OIOS 2015 (A/70/72), pág. 6.

### 3.2.4 La evaluación en los Programas y Fondos

La principal característica de los Programas y Fondos de NU es que su financiación proviene en gran medida de contribuciones voluntarias<sup>62</sup>. Su gran tamaño en cuanto a mandato y cobertura ha incidido en un desarrollo sólido de la función de evaluación con unidades de evaluación propia, políticas de evaluación y con un alto grado de independencia. En los Programas y Fondos, la principal unidad de intervención son los programas y proyectos, que se desarrollan en base a un ciclo programático. En funciones de evaluación sólidas, como en este caso, las evaluaciones también forman parte de dicho ciclo de programas, y como tal, las unidades de evaluación son las encargadas de elaborar un plan de evaluación generalmente bianual en base a unos criterios pre-establecidos, con el fin de incorporar las recomendaciones de la evaluación en el ciclo del programa. A continuación se enumeran los Fondos y Programas existentes:

- ✓ El Programa de las Naciones Unidas para el Desarrollo (PNUD)
- ✓ El Fondo de las Naciones Unidas para la Infancia (UNICEF)

<sup>62</sup> En algunos casos, como en el caso de UNICEF las contribuciones voluntarias son el 100%. En otros casos, como PNUD o PMA, las contribuciones voluntarias son también complementadas con contribuciones regulares de los países.

- ✓ El Programa Mundial de Alimentos (PMA)
- ✓ El Fondo de Población de las Naciones Unidas (FNUAP)
- ✓ El Programa de Naciones Unidas para el Medioambiente (PNUMA)
- ✓ El Programa de Naciones Unidas para los Asentamientos Humanos (ONU-Habitat)
- ✓ El Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR)

Dado su gran tamaño, algunos de estos programas y fondos combinan estructuras centralizadas y descentralizadas de evaluación dentro de la propia organización. De esta manera, las oficinas de evaluación central conducen evaluaciones de programa, temáticas y, en algunos casos, a nivel de país, a la vez que proporcionan cierto apoyo metodológico a las funciones de evaluación descentralizadas. Por su parte, las oficinas de país son las responsables de la gestión de las evaluaciones a nivel descentralizado. Generalmente, las oficinas de evaluación central de estos programas y fondos reportan a sus respectivas Juntas de Gobierno a través de la Dirección General, o directamente. PNUD y UNICEF son algunos ejemplos donde la evaluación esta principalmente descentralizada.

### **3.2.5 La evaluación en organismos especializados**

Se llaman organismos especializados a aquellas organizaciones autónomas que trabajan con las Naciones Unidas a través de los mecanismos de coordinación establecidos por el Consejo Económico y Social en el plano intergubernamental, y bajo acuerdos negociados. A continuación se enumera el grupo de organismos especializados del Sistema de NU:

- ✓ Organización Internacional del Trabajo (OIT)
- ✓ Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO)
- ✓ Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO)
- ✓ Organización Mundial de la Salud (OMS)
- ✓ Grupo del Banco Mundial
- ✓ Fondo Monetario Internacional (FMI)
- ✓ Organización de Aviación Civil Internacional (OACI)
- ✓ Organización Marítima Internacional (OMI)

- ✓ Unión Internacional de Telecomunicaciones (UIT)
- ✓ Unión Postal Universal (UPU)
- ✓ Organización Meteorológica Mundial (OMM)
- ✓ Organización Mundial de la Propiedad Intelectual (OMPI)
- ✓ Fondo Internacional de Desarrollo Agrícola (FIDA)
- ✓ Organización de las Naciones Unidas para el Desarrollo Industrial (ONUDI)
- ✓ Organización Mundial del Turismo (OMT)

Estas organizaciones presentan un alto grado de heterogeneidad en cuanto a la solidez de sus funciones de evaluación. En algunos casos como el FIDA, el Grupo del Banco Mundial, y la ONUDI tienen establecidas oficinas centrales de evaluación con alto nivel de independencia con respecto de los programas ejecutados, recursos propios, y políticas de evaluación propias. En algunos casos de alto nivel de independencia, las oficinas de evaluación reportan directamente a la Junta de Gobierno, y tienen directores/as de evaluación directamente elegidos y presupuestos aprobados por dicha Junta de Gobierno<sup>63</sup>.

### **3.2.6 La evaluación en organizaciones relacionadas**

Como en los organismos especializados, las organizaciones relacionadas son independientes. Se nutren de cuotas obligatorias y aportaciones voluntarias. A diferencia de los organismos especializados, informan a la Asamblea General o al Consejo de Seguridad.

- ✓ Comisión Preparatoria de la Organización del Tratado de prohibición completa de los ensayos nucleares (CTBTO)
- ✓ Organismo Internacional de la Energía Atómica (OIEA)
- ✓ Organización Mundial del Comercio (OMC)
- ✓ Organización para la prohibición de las Armas Químicas (OPAQ)

En las organizaciones relacionadas, la evaluación está establecida como función y, en mayor o menor medida, las actividades de evaluación son planificadas por unidades de evaluación que reportan directamente a la Dirección General de cada Organización. Se puede afirmar que la evaluación es una práctica más incipiente en estas organizaciones, en comparación con los organismos especializados, o con los programas y fondos que reportan a la

---

<sup>63</sup> Este es el caso de FMAM, organismo independiente pero administrativamente vinculado al Grupo del Banco Mundial.

Asamblea General. El carácter especialmente técnico de estas organizaciones y su vinculación menos directa con la consecución de resultados sociales y de desarrollo, hacen que, hasta la fecha, la práctica evaluativa en estas organizaciones no haya tenido un rol prioritario como sucede en los otros organismos.

Sin embargo, en los últimos años, se han producido importantes avances en la institucionalización de la función de evaluación también en estos organismos. Este es el caso de la OIEA que, en los últimos años, ha realizado importantes esfuerzos de re-alinear su política de evaluación a las Normas UNEG, y en el establecimiento de un sistema de seguimiento a las recomendaciones de evaluación que facilita su integración en la formulación del ciclo de programas.

### **3.3. El marco normativo sobre evaluación en NU**

#### **3.3.1 El rol de la evaluación en el ciclo programático**

Tal y como se ya se señalado, una de las unidades de intervención más frecuentes en el Sistema de NU es el programa o proyecto. Los programas y proyectos son aprobados de manera multianual por los diferentes órganos de gobierno, en lo que constituye el ciclo programático.

Las directrices legislativas que rigen la planificación, programación, supervisión y evaluación de las intervenciones de las NU fueron aprobadas por la Asamblea General en su resolución 37/234 del 21 de diciembre de 1982. Dicha resolución fue revisada en 1987 (ST/SGB/PPBME Rules 1) y en 2000 (ST/SGB/2000/8) a través del boletín del Secretario General: "*Reglamento y reglamentación detallada para la planificación de los programas, los aspectos de programas del presupuesto, la supervisión de la ejecución y los métodos de evaluación*", más conocido como PPBME por sus siglas en inglés<sup>64</sup>.

El boletín PPBME está compuesto de diferentes artículos en los cuales se explican aspectos referentes a la aplicabilidad de las normas, los instrumentos de planificación integrada, el proceso de planificación y de presupuestación, el plan de medio plazo y las actividades de monitoreo y evaluación. Como puntos importantes de este boletín destacan los siguientes:

- ✓ Todas las actividades de Naciones Unidas están sujetas a un proceso de planificación, monitoreo y evaluación, independientemente de la fuente de los fondos, de la regularidad, o voluntariedad de los mismos.
- ✓ Los instrumentos clave de gestión integrada son: planes de medio-plazo (2 bienios), presupuestos en base a programas, informes de seguimiento y monitoreo; e informes de evaluación.
- ✓ El ciclo de planificación, programación, presupuestación y evaluación forma parte integrante del proceso de gestión de NU. Los instrumentos

---

<sup>64</sup> Última revisión del 2000 accesible en <http://daccess-dds-ny.un.org/doc/UNDOC/GEN/N00/408/48/PDF/N0040848.pdf?OpenElement>

de gestión integrada sirven para asegurar la coordinación necesaria, y una utilización eficiente de los recursos.

- ✓ Siguiendo un enfoque estratégico, los planes de medio-plazo (4 años) deben servir para la formulación de programas y presupuestos bianuales, de los cuales se deriven los planes anuales. Estos planes de medio-plazo deben cubrir las actividades de presupuesto regular y todas aquellas que formen parte de partidas extraordinarias de fondos.
- ✓ El enfoque es programático, por lo cual, los programas y subprogramas son las principales unidades de análisis de los planes de medio-plazo, y otras actividades de monitoreo y evaluación. Puesto que los planes se basan en programas dinámicos, estos primeros son objeto de revisión y actualización a mitad de término (2 años).
- ✓ Los logros son medidos por el grado de obtención de los productos previstos en el presupuesto. Una vez terminado el bienio que abarca el presupuesto, el Secretario General informa a la Asamblea General, a través del Comité del Programa y la Coordinación, sobre la ejecución de los programas durante ese periodo.
- ✓ La evaluación tiene 2 objetivos: la determinación sistemática y objetiva de la pertinencia, eficiencia, efectividad y efectos de las actividades de las NU en relación con sus objetivos; y facilitar que la Secretaría y los Estados Miembros reflexionen de forma sistemática con miras a aumentar la efectividad de los programas principales de la Organización, modificando su contenido y, de ser necesario, revisando sus objetivos.
- ✓ Los resultados de los estudios de evaluación efectuados se reflejan en el diseño y ejecución de los programas y en las directrices políticas. Con este propósito se presenta a la Asamblea General, de manera paralela al texto del proyecto de plan de mediano plazo, un informe breve en el que se resumen las conclusiones del Secretario General acerca de todos los estudios de evaluación realizados dentro del programa de evaluación establecido.

El boletín ST/SGB/2000/8 es referencia clave de todos los esfuerzos impulsados por los Estados Miembros, a través de la Asamblea General, para armonizar la evaluación dentro del ciclo de proyecto, y otorgarle una importancia clave en el análisis de los logros obtenidos, la modificación de objetivos, y el establecimiento de prioridades programáticas a ser presupuestadas. A través de esta regulación, la evaluación es considerada un instrumento integrante del ciclo programático, con un rol claro de proporcionar información que guíe el análisis de los resultados obtenidos, y la aprobación de programas futuros.

### 3.3.2 Las Normas y Estándares UNEG

Además de la regulación PPBME, que como se ha mencionado en el apartado anterior es aplicable a todo el Sistema de las NU, también existe dentro del Sistema un foro de discusión sobre aspectos de evaluación denominado Grupo de Evaluación de las Naciones Unidas (*United Nations Evaluation Group*, UNEG). La primera reunión tuvo lugar en enero de 1984 (en aquel entonces se llamaba Grupo de trabajo interagencial de evaluación, *Inter Agency Working Group in Evaluation*) y desde entonces, el foro se ha reunido anualmente para debatir cuestiones de interés común sobre evaluación en las diferentes agencias del Sistema.

De esta manera, la UNEG se constituye como un grupo de trabajo que reúne periódicamente a los máximos responsables de las unidades de evaluación dentro de los organismos del Sistema de las NU y que ha trabajado activamente no solo en el ámbito dialéctico sino también como órgano generador de regulación clave en materia de evaluación. Así, en respuesta a las resoluciones de la Asamblea General, la UNEG ha contribuido desde sus orígenes a la profesionalización de la práctica evaluativa en el Sistema de las NU a través de la elaboración de normas, estándares y procedimientos. Actualmente, 46 organismos de las NU son miembros de UNEG y el PNUD ejerce el papel de Secretaría.

Como también se ha apuntado, la Asamblea General y la Secretaría General de las NU han realizado reiterados pronunciamientos sobre la necesidad de fortalecer la evaluación en las NU. En diciembre de 2004, la Asamblea General, a través de la resolución A/RES/59/250<sup>65</sup> solicitó a UNEG que hiciera un esfuerzo de armonización y simplificación de metodologías, normas y estándares de evaluación entre las diferentes agencias del Sistema. La profesionalización de las funciones de evaluación es considerada como una condición necesaria para el Sistema de las NU, tal y como ha sido señalado en las revisiones de la Política Integral Cuatrienal<sup>66</sup> (resoluciones GA 59/250 de 2004 y 62/208 de 2007).

Desde su establecimiento, las tareas de la UNEG han ido encaminadas a la promoción de la evaluación como principal herramienta de aprendizaje y rendición de cuentas dentro del Sistema de las NU. A continuación, se resumen las principales funciones que, con este propósito, ha desempeñado la UNEG:

- ✓ Promocionar la adopción y aplicación de una serie de normas y estándares como instrumento de armonización de procedimientos evaluativos en las NU.
- ✓ Fortalecer las capacidades técnicas y profesionales de las personas encargadas de la evaluación.
- ✓ Facilitar el intercambio de experiencias y aprendizaje en evaluación.

---

<sup>65</sup>A/C.2/59/L.63, 17 Diciembre 2004, para 69

<sup>66</sup> Hasta 2007 las revisiones de Política Integra tenían una periodicidad trienal.

- ✓ Promover iniciativas conjuntas entre organismos del Sistema de las NU en materia evaluativa.
- ✓ Promocionar el uso de la evaluación para el aprendizaje, la toma de decisiones y la rendición de cuentas.
- ✓ Facilitar el apoyo necesario a países miembros para desarrollar su capacidad evaluativa a nivel nacional en diferentes ámbitos, de forma que puedan evaluar sus propios programas;
- ✓ Facilitar el asociacionismo y el desarrollo de capacidades con otras redes de evaluación, incluyendo las redes regionales de evaluación OCDE/DAC y de otros organismos multilaterales.

Es importante subrayar que uno de los primeros y más relevantes aportes de la UNEG a la comunidad evaluativa de las NU ha sido el desarrollo y difusión en abril de 2005 de las "*Normas y Estándares para la Evaluación en el Sistema de las NU*". Para su elaboración, UNEG utilizó como un punto de referencia clave los principios de evaluación de la OCDE/DAC y los estándares nacionales de los países miembros de la OCDE.

*"Las normas UNEG buscan facilitar la colaboración, en materia de evaluación, entre todas las organizaciones del sistema, asegurando que las entidades de evaluación que forman parte de las NU se ciñan a principios básicos acordados. Las normas constituyen una referencia para fortalecer, profesionalizar y mejorar la calidad de la evaluación en todas las entidades del Sistema de las NU, incluyendo fondos, programas y agencias especializadas"* (Normas UNEG 2005, pág. 2).

Por su parte, los estándares UNEG sirven de complemento a las Normas UNEG, y han sido seleccionados a través de lo que se denominan '*buenas prácticas*' entre los miembros de las organizaciones que componen la UNEG. Así, los estándares "*han sido diseñados para guiar el establecimiento del marco institucional, la gestión de la función de evaluación y la realización y el uso de evaluaciones. Representan, asimismo, un referente para las competencias de los profesionales y la ética de trabajo en materia de evaluación, y la intención es que sean aplicados dentro de cada organización según resulte apropiado*" (Estándares UNEG 2005, pág. 2). Los estándares UNEG se centran en 4 áreas fundamentales: i) marco institucional y gestión de la función de evaluación; ii) competencias y aspectos éticos; iii) realización de evaluaciones (estándares que cubren desde el diseño hasta el seguimiento de las recomendaciones; y iv) los informes de evaluación.

En relación a las normas UNEG, debido a su relevancia, a continuación se analizan los principales temas tratados en cada una de las ellas:

#### Norma 1 - Definición

Proporciona una definición de lo que es evaluación. "*Evaluación es una valoración, lo más sistemática e imparcial posible, de una actividad,*

*proyecto, programa, estrategia, política, tópico, sector, área operativa, desempeño institucional, etc.”*

Como se puede apreciar, se trata de una adaptación de la definición propuesta por DAC/OCDE, aunque más elaborada en cuanto a su ámbito de acción. Además, especifica los principales objetivos de la práctica evaluativa en torno a dos pilares fundamentales: rendición de cuentas (*accountability*) y construcción de conocimiento (*knowledge building*). También destaca que la evaluación forma parte del ciclo proyecto, contribuyendo a su vez a la gestión por resultados y alimentando la toma de decisiones. A su vez, puntualiza los diferentes tipos de valoración que pueden hacerse en relación al objeto evaluado, diferenciando así entre evaluación de ‘seguimiento’, ‘investigación’ y ‘auditoría’ entre otros. Finalmente, señala los diferentes tipos de evaluación posibles en función de: i) quién realiza la evaluación (personal externo o interno), ii) el enfoque (sumativo o formativo), y iii) el momento en que se realiza la evaluación (a medio término o ex post).

## Norma 2 - Responsabilidad frente a la evaluación

Define cuáles son los órganos de gobierno y/o de dirección de las organizaciones que tienen la responsabilidad de ubicar la función de evaluación dentro del organigrama de la institución así como de garantizar un entorno favorable para el desempeño de la práctica evaluativa. Un entorno favorable implica asegurarse de que existen recursos adecuados asignados para la función de evaluación y que las evaluaciones se puedan llevar a cabo de manera imparcial y autónoma.

Además, señala que dichos órganos de gobierno y/o de dirección deben designar a una persona responsable de la evaluación competente, asegurar que la evaluación contribuya a la toma de decisiones así como garantizar que se establezca un sistema adecuado para la planificación de evaluaciones y, a su vez, monitorear el grado de implementación de las recomendaciones que emanen de los informes de evaluación.

Finalmente los órganos de gobierno y/o de dirección de las organizaciones son los responsables últimos de asegurar que exista un repositorio de las evaluaciones realizadas y un mecanismo para disseminar y difundir de forma amplia las evaluaciones.

## Norma 3 - Política de evaluación

Determina que cada organización debe desarrollar una política específica en materia de práctica evaluativa. Dicha política debe proveer una clara explicación del concepto, rol y uso de la evaluación en la organización, considerando en todo momento el contexto institucional. Igualmente, la política de evaluación debe definir roles y responsabilidades en materia evaluativa; explicar la forma y procedimientos utilizados para planear, gestionar y presupuestar las

evaluaciones; y elaborar un claro plan de acción para diseminar los resultados obtenidos a partir de la práctica evaluativa.

La Norma 3 se amplía en el Estándar 1.2 donde se especifica que las organizaciones del Sistema de las NU deben desarrollar una política de evaluación y actualizarla regularmente teniendo en cuenta las Normas y Estándares para Evaluación.

Adicionalmente, el Estándar 1.2 destaca la importancia de obtener la aprobación de la política de evaluación al más alto nivel posible dentro de las instituciones, indicando que la política de evaluación debe ser aprobada por la Junta de Gobierno de las agencias y/o por su máximo responsable; y debe estar en línea con las Normas y Estándares UNEG así como con los objetivos y estrategias propios de cada organismo.

#### Norma 4 - Intencionalidad

Introduce el concepto de utilidad como parte inherente de la práctica evaluativa enfatizando que las evaluaciones deben ser planificadas e implementadas con el propósito último de que sus hallazgos sean utilizados. Para ello, una planificación adecuada y sistemática se considera clave. En cuanto a cómo elaborar el plan de evaluación, la Norma 4 presenta la posibilidad de realizar una selección de los temas de evaluación en base a un criterio preestablecido.

Asimismo, esta norma subraya que la razón de ser de una evaluación y de las decisiones basadas en ella debe estar clara desde el principio. El alcance y el diseño de la evaluación deben formularse con la intención de generar productos pertinentes y oportunos que respondan a las necesidades de los usuarios previstos.

#### Norma 5 - Imparcialidad

Define las implicaciones de ser imparcial (ausencia de sesgos en el proceso, metodología y presentación de conclusiones) al mismo tiempo que se consideran todos los puntos de vista de las personas, grupos interesados. Determina, además, que, en caso de divergencia de los mismos, estas diferencias deben reflejarse en la evaluación.

También, menciona los beneficios de ser imparcial, entre los que se encuentra el incremento de la credibilidad y legitimación de la práctica evaluativa. Dicha imparcialidad, para ser completa, debe comprender todas las fases del proceso de evaluación, desde la planificación de la misma hasta la formulación de recomendaciones.

#### Norma 6 - Independencia

El concepto de independencia está relacionado con la ubicación de la función de evaluación en la estructura de la organización. Así, la Norma 6 considera que la función de evaluación debe ser estructuralmente

independiente de las funciones de gestión operacional y de adopción de las decisiones en la organización con el fin de que esté libre de influencias y tenga plena autoridad para presentar informes directamente en los niveles apropiados de toma de decisiones.

La norma enfatiza que es necesario que los equipos de evaluación sean a su vez independientes en el sentido de que no hayan participado en la planificación o implementación del programa objeto de evaluación. Es importante señalar que, en este contexto, independiente no implica necesariamente ser externo a la organización, sino fuera de la línea de mando de los órganos responsables del programa evaluado.

#### Norma 7 - Evaluabilidad

La norma 7 recomienda llevar a cabo un ejercicio de 'evaluabilidad' antes de involucrarse en la realización de evaluaciones que requieran de una gran cantidad de tiempo y recursos. Entre los criterios a tener en cuenta en dicho ejercicio de evaluabilidad se incluyen: la claridad del objeto evaluado, la existencia de indicadores suficientes, de fuentes de información fiables y de garantías de que el proceso de evaluación no va a estar sometido a ninguna presión que amenace su imparcialidad.

#### Norma 8 - Calidad de la Evaluación

En general, la Norma 8 enfatiza que las evaluaciones deben responder a las normas mínimas de calidad que se definan por la unidad encargada de la evaluación en la organización. En este sentido, el plan de evaluación debe, a su vez, ser práctico y eficaz en cuanto a costes. Para garantizar que la información que se genere sea exacta y fiable, el diseño de la evaluación así como la recogida y el análisis de los datos deben reflejar unos estándares de calidad mínimos teniendo debidamente en cuenta cualquier circunstancia o limitación existente.

#### Norma 9 - Competencias en Evaluación

Norma que alude a la necesidad de contar con criterios formales para la selección del personal evaluador. Asimismo, se destaca la importancia de que la persona responsable de la función de evaluación en la organización disponga de las competencias adecuadas para la gestión y realización de estudios de evaluación.

#### Norma 10 - Transparencia y consulta

Destaca la necesidad de que para lograr la credibilidad así como la utilidad de la evaluación es importante consultar a las partes interesadas, además de hacer públicos los informes y diseminar los resultados obtenidos. Esta forma de actuar facilita la construcción de consensos y una mayor responsabilidad a la hora de interpretar los hallazgos y de hacer frente a las recomendaciones que emanan del informe de evaluación.

## Norma 11 - Ética de la evaluación

Enfatiza que la evaluación no debe reflejar intereses personales ni sectoriales. Las personas encargadas de la evaluación deben poseer integridad profesional y deben, asimismo, respetar el derecho de las instituciones y de los particulares de confidencialidad y garantizar la verificación en relación a las afirmaciones que se les atribuyen.

## Norma 12 - Seguimiento de una evaluación

Hace referencia a los pasos a tener en cuenta una vez el informe de evaluación ha sido emitido. Primeramente, son los equipos responsables del programa evaluado quienes deben dar sus comentarios, incluyendo un plan de acción para las recomendaciones aceptadas. A continuación, dicho plan de acción deberá ser objeto de una revisión periódica para conocer el progreso en la aplicación de las recomendaciones. La función de evaluación es responsable de recopilar el estado de la aplicación de las recomendaciones de evaluación y de presentarlo a los órganos de gobierno y/o máximo responsable de la organización.

## Norma 13 - Contribución a la creación de conocimiento

Especifica los dos aspectos principales para contribuir a la creación de conocimiento y, por tanto, a la mejora de la propia marcha de la organización. Por una parte, el lenguaje y la manera en que se presenten los hallazgos de la evaluación deben ser acorde al público al que va destinado, de tal manera que se comprenda adecuadamente su contenido. Por otra parte, la evaluación debe ser accesible a su público objetivo. Para ello, se considera necesaria la aplicación de una política clara de difusión de las evaluaciones.

La importancia de las normas y estándares UNEG reside en el extenso acuerdo dentro de las instituciones del Sistema de las NU sobre la necesidad de estas normas y estándares, de manera que se conviertan en un referente general de la práctica evaluativa dentro del Sistema NU.

### **3.4. La capacidad evaluativa en las NU**

#### **3.4.1 Principios rectores de la capacidad evaluativa**

La capacidad evaluativa ha adquirido un progresivo protagonismo dentro del Sistema de las NU en los últimos años. Así, el PNUD define capacidad evaluativa como “*la habilidad de los individuos, instituciones y sociedades de realizar funciones, solucionar problemas y definir y lograr objetivos de una manera sostenible*” (UNDP 2008, pág. 4). El desarrollo de capacidades se entiende como un proceso continuo a través del cual dichas habilidades de

gestionar objetivos y estrategias son obtenidas, fortalecidas y, si se requiere, modificadas.

A su vez, la UNEG enfatiza que la capacidad evaluativa hace referencia a la capacidad de diseñar, gestionar y utilizar evaluaciones en un organismo determinado. En este sentido, el desarrollo de la capacidad evaluativa puede describirse como el proceso a través del cual las habilidades para diseñar, gestionar y utilizar las evaluaciones son adquiridas, fortalecidas, adaptadas y sostenibles a través del tiempo.

El elemento clave a través del cual se canalizan las estrategias de desarrollo de capacidad evaluativa en las NU es mediante del fortalecimiento de las funciones de evaluación. Así, en el año 2008, en uno de los grupos de trabajo de la UNEG sobre revisión por pares (*peer reviews*) de las funciones de evaluación de los diferentes organismos dentro de las NU, se identificaron los siguientes principios centrales para lograr una mayor capacidad evaluativa a partir del fortalecimiento de la función de evaluación (DAC/UNEG/2007):

- ✓ Independencia: La práctica evaluativa ha de estar libre de influencias organizacionales o personales que puedan sesgar el proceso de evaluación, sus hallazgos, sus conclusiones y/o sus recomendaciones. Esto implica que las evaluaciones deben ser gestionadas desde unidades que están fuera del control de los equipos responsables del diseño y de la implementación de los programas evaluados. La independencia de la función de evaluación es considerada un requisito fundamental para su credibilidad, validez y, por tanto, utilidad. De esta forma, es preciso prestar atención a la aplicación de medidas y procedimientos que garanticen la objetividad e imparcialidad necesaria de los procesos evaluativos. Estos indicadores de independencia están cubiertos por las Normas N6.1-N6.5.
- ✓ Credibilidad: Los contenidos que emanen de la práctica evaluativa han de ser confiables y verdaderos. La credibilidad en evaluación está relacionada con diversos factores como, por ejemplo, la transparencia del proceso de evaluación, la precisión del contenido del informe y la percepción de que los equipos de evaluación son neutrales y competentes. La credibilidad exige que las evaluaciones informen sobre los éxitos y los fracasos de los programas y, para ello, debe existir una amplia participación de los actores clave, incluyendo países beneficiarios, de cara a promover la necesaria credibilidad y compromiso con los resultados de la evaluación. Las organizaciones deben dedicar sus esfuerzos a promover así el empoderamiento de la práctica evaluativa. Los indicadores de credibilidad están reflejados en las Normas UNEG N5.1-N5.3, N8.1-8.2, N9.1-N9.3 y N11.1-N11.5.
- ✓ Utilidad: El propósito de la práctica evaluativa es que se utilicen sus resultados de forma que éstos informen e influyan en la toma de decisiones de un organismo determinado. Este principio trata de enfatizar que las evaluaciones tienen el fin de generar cambios que permitan a las instituciones y, en el caso de las , también a los gobiernos

lograr sus objetivos de manera eficiente. Para ello, las evaluaciones deben reflejar las necesidades de los actores clave. La responsabilidad en la utilidad de la evaluación es compartida entre los equipos de evaluación y el resto de los actores clave, incluyendo a los máximos responsables de los organismos en cuestión así como a los países miembros a través de su participación en los órganos de gobierno de dichos organismos. Las normas UNEG que cubren el principio de utilidad son principalmente N1.3, N2.6, N8.2, N10.1-10.2, y N12.1-12.3.

La independencia, la credibilidad y la utilidad de las evaluaciones son principios comúnmente aceptados dentro del Sistema de las NU así como la interrelación existente entre dichos principios. Por ejemplo, una evaluación sesgada inclinada a reflejar la perspectiva de un colectivo (y, por tanto, no imparcial), probablemente no sea muy creíble para otros colectivos. En el momento en que una evaluación deja de ser creíble, su utilidad automáticamente disminuye.

En la práctica, las dificultades inherentes a la necesidad de equilibrio entre los principios centrales (independencia, credibilidad y utilidad) son evidentes. Las presiones institucionales y de los donantes en la elección del plan de evaluación y de los temas a evaluar, la dependencia de fondos e, incluso, las propias personas que realizan las evaluaciones pueden poner en peligro la independencia de la práctica evaluativa y tener repercusiones muy negativas sobre la credibilidad de dichas evaluaciones. Así, por ejemplo, las evaluaciones de donde emanan resultados positivos con abundantes buenas prácticas suelen ser mejor acogidas por donantes y personal interno de la organización, que aquellas evaluaciones donde se ponen de manifiesto hallazgos relativos a serios fracasos en la implementación del objeto evaluado. Esto, en ocasiones, se puede convertir en un serio limitante de la correcta y adecuada práctica evaluativa al generar presiones sobre los distintos actores que participan en ella dirigidas a reflejar únicamente aspectos positivos, lo cual en última instancia afecta gravemente a su utilidad.

Así, el balance entre los tres principios es clave para el desarrollo de la capacidad evaluativa en una institución. Sin embargo, es importante señalar que la responsabilidad de poner en práctica medidas para fortalecer el equilibrio entre estos principios es compartida entre diferentes actores dentro de un organismo. Las unidades de evaluación son responsables de producir evaluaciones creíbles y de calidad, pero la utilidad de las mismas y la independencia del proceso de evaluación son aspectos que no están enteramente bajo el control del equipo evaluador al depender del apoyo institucional y del compromiso de los responsables de la toma de decisión con la consolidación de un sistema de evaluación independiente y profesional.

### **3.4.2 Instrumentos de análisis de la función de evaluación**

En el Sistema de las NU existen principalmente dos tipos de análisis que miden la calidad del desempeño de la función de evaluación en los organismos que lo componen y, en concreto, la capacidad de incidencia de los hallazgos evaluativos en la rendición de cuentas y en el aprendizaje organizacional. El primero es el que comúnmente se denomina revisión por pares (*peer review*).

La revisión por pares es realizada por paneles de evaluadores/as independientes de acuerdo con los procedimientos especificados por la UNEG. El segundo son los informes bianuales realizados por OIOS. A continuación, se describen los aspectos más relevantes de ambos procedimientos.

### **Revisión por pares**

En el marco de la UNEG, se realiza lo que se denomina revisión por pares o *peer review*, que consiste en un análisis exhaustivo realizado por profesionales de la evaluación (*peers*) sobre el estado de las funciones de evaluación de las organizaciones de las NU aplicando los tres principios descritos anteriormente (independencia, utilidad y credibilidad). Este análisis aporta elementos muy relevantes sobre la capacidad evaluativa de las organizaciones puesto que analiza, con detalle y con rigor metodológico, la existencia y puesta en marcha en un organismo determinado de normas y procedimientos que faciliten la utilidad, credibilidad e independencia de la práctica evaluativa.

La principal pregunta a la que un panel compuesto por evaluadores/as profesionales de acuerdo con las Normas y Estándares UNEG (UNEG/REF/2011/1) trata de responder a través de este procedimiento es hasta qué punto la política de evaluación de una organización así como su función evaluativa y el conocimiento que genera son independientes, creíbles y útiles para el aprendizaje y la rendición de cuentas.

El proceso de revisión por pares responde así a un proceso de valoración sistemática del rendimiento de la función de evaluación de una organización desde una perspectiva profesional y externa. El objetivo último de este procedimiento es asistir a la organización objeto de estudio a mejorar su desarrollo de políticas en materia evaluativa y a alinearse con los estándares de calidad y principios generalmente establecidos y aceptados. No se trata de un procedimiento que derive en acciones punitivas o sancionadoras, sino que, por el contrario, persigue incentivar y estimular a las organizaciones a efectuar aquellos cambios que les permitan lograr sus objetivos mediante el cumplimiento de estándares evaluativos (UNEG/REF/2011/1). En concreto, los propósitos inherentes a los ejercicios de revisión por pares son:

- ✓ Proporcionar una valoración independiente y profesional sobre el grado de adopción de las Normas y Estándares UNEG en las organizaciones. Dicha valoración puede conllevar a la puesta en marcha de medidas para mejorar la política y práctica evaluativas.
- ✓ Generar un mayor conocimiento, confianza y utilidad de la práctica evaluativa por parte de la gerencia de un organismo determinado, así como de sus órganos de gobierno u otros usuarios interesados. Un mejor entendimiento de la calidad de la práctica evaluativa existente en una organización determinada puede poner de relieve, de forma efectiva, la necesidad de emprender acciones de mejora en este campo que se traduzcan en una mayor integración de la función de evaluación

en el sistema de información y de gestión por resultados, así como en una mejora general de la práctica evaluativa.

- ✓ Proporcionar una herramienta consistente para 'evaluar a los evaluadores/as' a través del trabajo de profesionales en el área evaluativa guiados por las Normas y Estándares UNEG y respetando, en todo momento, el grado necesario de independencia que esta tarea requiere.
  
- ✓ Compartir buenas prácticas y experiencias de cara a contribuir al desarrollo de una mayor capacidad evaluativa que incremente la confianza en los sistemas de evaluación de las NU y reduzca la necesidad de obtener valoraciones externas en este terreno.

La principal audiencia de estos ejercicios de revisión por pares es el personal responsable de la toma de decisión en las organizaciones, como la gerencia, los órganos de gobierno y/o el personal responsable de la propia función de evaluación. También son destinatarios de este tipo de ejercicios otros actores clave en el entorno de la cooperación internacional incluyendo a usuarios finales procedentes de los países miembros.

El principal foco de atención de la revisión por pares es la función de evaluación de la organización. Por ello, la política de evaluación y la unidad de evaluación son los primeros elementos de análisis en este procedimiento, para así poder desvelar la forma en que se realiza la evaluación en la organización objeto de estudio. Después, se procede a analizar las prácticas evaluativas que emanan de las directrices de la política de evaluación, con especial énfasis en los procesos de planificación y presupuestación de la evaluación así como en el sistema de seguimiento de las recomendaciones. El análisis se centra también en la idoneidad de los canales de comunicación y coordinación entre varios niveles dentro de la organización como, por ejemplo, entre los órganos de gobierno, las unidades técnicas desde donde se diseñan e implementan los programas y la propia unidad de evaluación.

Existen procedimientos establecidos para garantizar el grado de experiencia e independencia del panel de revisión así como una descripción de los diferentes roles que pueden ejercer los y las profesionales integrantes del panel. En cuanto a la estructura del proceso en sí mismo, éste es similar al de cualquier ejercicio de evaluación incluyendo así una fase de planificación en la que se preparan los términos de referencia, se establecen las preguntas claves y se desgrana la metodología; una fase de recogida de información a través de cuestionarios, entrevistas, visitas de campo, grupos focales etc. incluyendo el

análisis y revisión documental; y culmina con la fase de preparación y diseminación del informe final de evaluación<sup>67</sup>.

En cuanto a su presupuesto, las revisiones por pares tienen un coste estimado de entre 50.000 a 100.000 dólares, distribuidos entre 50 y 100 días de trabajo. La duración media de un proyecto de revisión por pares oscila entre los 6 y 9 meses (UNEG/REF/2011/1). Hasta la fecha, existen un total de diez organismos que se han sometido al proceso de revisión por pares sobre el grado de adopción de las Normas y Estándares UNEG en las organizaciones de las NU<sup>68</sup>. Los organismos y las fechas de publicación de los informes son:

- ✓ ONU Habitat, Agosto 2012
- ✓ Programa de las Naciones Unidas para el Desarrollo (PNUD), Enero 2006 y Enero 2013
- ✓ Oficina de Servicios de Supervisión Interna (OIOS), Marzo 2009
- ✓ Programa de las Naciones Unidas para el Medioambiente (PNUMA), Agosto 2012
- ✓ Fondo para la Infancia de las Naciones Unidas (UNICEF), Mayo 2006
- ✓ Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), Septiembre 2012
- ✓ Organización de las Naciones Unidas para el Desarrollo Industrial (ONUDI), Agosto 2010
- ✓ Fondo Mundial para el Medioambiente (FMAM), Julio 2009, y Octubre 2014;
- ✓ Programa Mundial de Alimentos (PMA), Noviembre 2007 y Octubre 2014;
- ✓ ONU Mujeres, Septiembre 2014

### **Informe para el fortalecimiento del rol de la evaluación**

OIOS presenta de manera bianual a la Asamblea General a través del Comité para el Programa y la Coordinación<sup>69</sup> y siguiendo el Reglamento ST/SGB/200/8, concretamente la norma 7.4, un informe de gran relevancia en materia evaluativa que se denomina "*Fortalecer el rol de evaluación y la*

---

<sup>67</sup>A pesar de que el proceso y la metodología pueda ser similar, la revisión por pares no es considerada una evaluación completa en sentido estricto. En principio es menos exhaustiva, a pesar de que se le exige una metodología rigurosa.

<sup>68</sup>Para más información, ver: <http://www.uneval.org/document/peer-review-reports>

<sup>69</sup> Este comité está establecido por las Resoluciones del Consejo Económico y Social 31/93 y 2008 (LX).

*aplicación de los hallazgos evaluativos en el diseño de programas, asistencia y las directivas políticas".*

El propósito de este informe es describir y analizar el estado de la función de la evaluación en los organismos que forman parte de la Secretaría de las NU, medir los avances en cuanto a capacidad evaluativa y destacar las principales dificultades existentes para una práctica evaluativa eficiente y sostenible. Más específicamente, el informe se enfoca en: i) el análisis de la capacidad, calidad y utilidad de la función de evaluación en las entidades de la Secretaría ; y ii) el análisis de los resultados clave y conclusiones contenidas en los informes de evaluación finalizados en el bienio objeto de estudio. Este informe basa su análisis de la capacidad evaluativa en cinco criterios: recursos, estructura, procesos, cultura y competencias.

El análisis se centra en las organizaciones que reportan directamente a la Secretaría, y que, por tanto, están bajo la supervisión general de OIOS. La metodología aplicada es variada e incluye los siguientes elementos:

- ✓ Análisis de los informes de evaluación finalizados en el bienio reportado con el fin de valorar su calidad, resultados clave y conclusiones.
- ✓ Cuestionarios on-line realizados por las instituciones a través de puntos focales previamente designados en los organismos.
- ✓ Entrevistas a los puntos focales de los organismos.
- ✓ Valoración de las políticas de evaluación disponibles en las instituciones<sup>70</sup>.
- ✓ Valoración de los presupuestos de evaluación.
- ✓ Talleres con los puntos focales para discutir los resultados obtenidos.

Las conclusiones del último informe publicado (Marzo 2015), en relación al bienio 2012-2013 son contundentes. De acuerdo al informe, la Asamblea General, en su resolución 67/226, enfatizó la importancia de que los organismos dispongan de funciones de evaluación independientes, creíbles y útiles, con recursos suficientes y una cultura de evaluación que promueva el uso de los hallazgos y las recomendaciones, que emanen del proceso evaluativo, en el desarrollo de sus políticas y en la mejora de la marcha de los organismos. De acuerdo con el informe, este objetivo no ha sido conseguido en la propia Secretaría, donde la evaluación, a pesar de ciertos avances con respecto a bienios anteriores, tiene aún que convertirse en una función sólida e integrada en el trabajo programático.

El informe, a su vez, destaca que la práctica evaluativa ha de poder nutrir la toma de decisiones estratégicas y guiar el futuro de las organizaciones,

---

<sup>70</sup>No siempre todas las instituciones tienen política de evaluación. Por ejemplo, en el último informe de Marzo 2015 (A/70/72), 22 de las 33 instituciones cubiertas en el estudio tenían políticas de evaluación.

siempre que se se realice de manera independiente y creíble. En este sentido, el informe puntualiza que, hoy en día, la evaluación es un recurso subutilizado con un uso limitado que se refleja en la falta de evidencia sobre la influencia de la práctica evaluativa en la toma de decisiones (A/70/72).

#### **4. Marco de análisis de la capacidad evaluativa**

El objetivo de esta tesis es profundizar en el análisis de las medidas dirigidas al desarrollo de la capacidad evaluativa y sus factores condicionantes. Así, a partir de la búsqueda de evidencia empírica, la investigación persigue revelar los factores más determinantes para el fortalecimiento de la capacidad evaluativa y que conducen y favorecen un mayor uso de los hallazgos evaluativos, así como facilitan la integración formal de la práctica evaluativa en la estructura organizacional.

En la primera parte (capítulos 1 y 2), se ha analizado el marco teórico sobre la práctica evaluativa y se han descrito los modelos conceptuales existentes para el desarrollo de capacidad evaluativa. A pesar de que la evidencia empírica que sostiene los modelos conceptuales es todavía limitada, éstos convergen en la identificación de una serie de factores que, en principio, provocan un impacto favorable en la capacidad evaluativa. Estos factores son: la capacidad de aprendizaje, el grado de institucionalización de la función evaluativa y el marco normativo que la respalda.

Seguidamente en el capítulo tres se ha puesto de manifiesto que la función de evaluación en el Sistema de las NU, aunque es un área de importancia e indudable interés, aún requiere de un mayor fortalecimiento en diferentes frentes al ofrecer todavía impactos limitados tanto en relación con la calidad de las evaluaciones como en su uso en los procesos de toma de decisiones.

Sobre la base de los capítulos anteriores y de los análisis que comprenden, este capítulo cuatro se centra en la proposición de un marco y un método de análisis que permita medir el grado de influencia de los factores que favorecen el desarrollo de la capacidad evaluativa identificados en el marco teórico, para poder proceder así a la verificación de nuestra hipótesis. Para ello, primero se procede a la sistematización de los principales factores del marco teórico (apartado 4.1), después se continúa con la justificación del Sistema de NU como contexto válido para contrastar de la hipótesis (apartado 4.2), y, finalmente, se describe la metodología propuesta (apartado 4.3), y se detalla más información sobre las entidades objeto de estudio (apartado 4.4 y Anexo 2).

##### **4.1. Condiciones de la capacidad evaluativa**

El marco teórico que analiza la práctica evaluativa es amplio. Tal y como ya se ha analizado en el capítulo 2, éste cubre la evaluación desde diversos ángulos, ofreciendo varios modelos que explican, a la vez, que predicen los resultados que se obtendrán dentro de las organizaciones si éstas ponen en marcha una serie de estrategias para el desarrollo de capacidad evaluativa. En este sentido, se puede concluir que todos los modelos teóricos revisados convergen alrededor de las siguientes ideas clave:

En primer lugar, se considera comúnmente aceptada la definición propuesta por Stockdill et al. 2002 sobre capacidad evaluativa:

*“La capacidad evaluativa hace referencia a un sistema de acciones intencionales y dependientes de un contexto orientadas a posibilitar de manera sostenible que las evaluaciones sean de alta calidad y cuya apropiada utilización se produzca como algo regular y ordinario dentro de las organizaciones” (Stockdill et al. 2002, pág. 8).*

Un aspecto clave a destacar de esta definición es el **carácter intencional** de la capacidad evaluativa dado que la capacidad evaluativa persigue que se genere el mayor grado posible de calidad y utilidad evaluativa dentro de las organizaciones. Asimismo, hay que destacar el **componente contextual** en el cual se desempeña la práctica evaluativa, componente que, como veremos más adelante, puede desempeñar un papel crucial en las evaluaciones.

A partir de la definición anterior, los modelos existentes de desarrollo de capacidad evaluativa tratan de desvelar los factores que facilitan el logro de los objetivos inherentes a la capacidad evaluativa (mejora de la calidad y el uso sostenible de la evaluación). Los modelos también buscan asegurar la puesta en marcha de estrategias efectivas para que se produzcan dichos factores facilitadores de capacidad evaluativa. En particular, los tres modelos más relevantes, Preskill y Boyle (2008), Suarez-Balcázar et al. (2010) y Labin et al. (2012) presentan grandes similitudes en relación a dichos factores.

Así, a pesar de que la terminología utilizada en los modelos puede presentar ciertas diferencias, lo cierto es que se identifican tres factores fundamentales, también denominados condiciones, que emergen de los tres modelos:

- i) el grado de institucionalización de la función de evaluación;
- ii) la existencia de políticas y prácticas evaluativas, y
- iii) la capacidad de aprendizaje evaluativo en las instituciones.

#### **4.1.1 Institucionalización de la función de evaluación**

##### **Independencia estructural**

Desde las instancias gerenciales de una organización se establecen las políticas correspondientes orientadas a definir los roles y estructura que las distintas funciones técnicas y administrativas tienen dentro de la organización, dando así lugar a diferentes formas de institucionalización de funciones. En el caso de la práctica evaluativa, la manera más habitual en que su institucionalización se produce es mediante la creación de un equipo de evaluación que se responsabilice tanto de la coordinación de las evaluaciones, como del desarrollo de la capacidad evaluativa necesaria para lograr que las evaluaciones realizadas sean efectivas (King 2002; Volkov y King 2007). La independencia dentro de la organización de la que goce dicho equipo o unidad determinará, en gran medida, el grado en que las evaluaciones serán llevadas a cabo con imparcialidad y objetividad y, por consiguiente, serán menos vulnerables a presiones tanto internas como externas.

Las Normas y Estándares UNEG han descrito de manera precisa los procedimientos para garantizar la debida estructura e institucionalización de la

función de evaluación en el Sistema de NU. En este sentido, las Normas 6.1 y 6.2 proponen la consolidación de unidades de evaluación que se encuentren fuera de la línea de mando de la gerencia técnica y que puedan reportar directamente al órgano competente de toma de decisiones en aras de garantizar su independencia e imparcialidad. A su vez, las Normas 2.1-2.4, 6.2-6.5 velan por el establecimiento de procedimientos que no solo garanticen la imparcialidad de los procesos sino que, a su vez, faciliten un ambiente adecuado para la investigación evaluativa, libre de presiones internas o externas.

La interconexión de los principios de independencia, credibilidad y utilidad en la función de evaluación revela que una función de evaluación independiente facilita que las instancias responsables de la toma de decisiones demanden evaluaciones imparciales que puedan nutrir de información veraz, objetiva y confiable sus debates y decisiones, lo cual incrementa de forma sustancial la utilidad de la función de evaluación (Heider 2011).

### **Recursos financieros**

Se considera que la función de evaluación esta institucionalizada cuando, además de operar a través de una estructura independiente, goza de recursos financieros para poder llevar a cabo su actividad programática.

Stockdill et al. (2002) identifican, en particular, que el personal, el espacio, el tiempo, el equipo e incluso los sistemas de documentación requieren de recursos financieros, de forma que su existencia o ausencia tienen un claro impacto en la capacidad evaluativa.

Las Normas y Estándares UNEG también hacen eco de la importancia de disponer de un presupuesto adecuado, transparente y suficiente para cubrir el plan de evaluación. En concreto, las Normas 2.3, 4.1 y 4.2 promueven que: i) la unidad de evaluación cuente con recursos necesarios para poder llevar a cabo su trabajo con efectividad e independencia; ii) que exista un presupuesto alineado con el plan de evaluación de tal manera que la planificación de evaluaciones se sustente en presupuestos confirmados y transparentes y no en asignaciones extra presupuestarias.

En el ámbito de la cooperación al desarrollo, las organizaciones que tienen un mayor contacto y acceso a diferentes grupos clave (donantes, órganos de gobierno, usuarios finales), siempre que éstos se encuentren concienciados y comprometidos con la evaluación, tendrán más oportunidades de asignar recursos adecuados a la función de evaluación (Suarez-Balcázar et al. 2010).

### **Recursos humanos**

La gestión del equipo de evaluación es un factor que tiene especial importancia en el grado de institucionalización de la función de evaluación. La credibilidad de la información evaluativa pasa por garantizar que las evaluaciones se lleven a cabo de manera imparcial y autónoma. En este sentido, los y las profesionales de la evaluación deben contar con la libertad necesaria para

llevar adelante su trabajo sin repercusiones para su evolución profesional (Normas UNEG 2.3, 2.4).

Además de la existencia de procedimientos que garanticen la imparcialidad y autonomía de los y las evaluadores/as, es fundamental contar con la debida competencia profesional. Las Normas UNEG exigen la garantía de que la función de evaluación sea ejercida por profesionales altamente competentes (Norma UNEG 2.5).

Las unidades de evaluación se enfrentan al desafío que supone combinar el conocimiento técnico del área temática a evaluar, con el conocimiento metodológico necesario para garantizar un debido proceso evaluativo en base a estándares aceptados. Puesto que los recursos financieros de la unidad de evaluación son limitados y no siempre se puede contar con la experiencia profesional suficiente en todas las áreas temáticas de la organización, las unidades de evaluación suelen estar compuestas por personal (*staff*) de la organización combinado con otros colaboradores externos (consultores/as), que proporcionan experiencia y conocimientos técnicos en las áreas particulares donde se enmarcan los proyectos y programas a evaluar.

En este sentido, las normas UNEG puntualizan la importancia del rol de gestión y supervisión del personal de evaluación (*staff*) (Norma 9.3), y de ello emana la necesidad de que exista un balance adecuado entre personal y consultores/as externos, así como una combinación apropiada de puestos de mayor experiencia con aquellos de menor experiencia, para lograr un equipo de evaluación competente y equilibrado en cuanto a conocimiento técnico y evaluativo se refiere.

#### **4.1.2 Políticas y prácticas evaluativas**

Prácticamente todas las aportaciones teóricas coinciden en que, al grado de institucionalización de la función de evaluación, se le debe agregar otro factor fundamental de cara a determinar en qué medida se establece la capacidad evaluativa adecuada dentro de una organización. Se trata, en este caso, de las políticas y prácticas evaluativas.

##### **La política de evaluación**

Heider (2007, pág. 86) define política en general como “*el curso de acción específico que permite la consecución de objetivos determinados*”. Las políticas, así, se dirigen a influenciar comportamientos dentro de las organizaciones puesto que tratan de prescribir un tipo de conducta determinada que, en el caso de la evaluación, se centraría en lograr que la práctica evaluativa sirva como herramienta útil para la rendición de cuentas, el aprendizaje y la generación de conocimiento.

De esta forma, en el ámbito de las NU, las Normas y Estándares UNEG destacan la necesidad de establecer políticas de evaluación dentro de cada organización. En concreto, dicho requisito está especificado en la Norma 3. A su vez, la División de Inspección y Evaluación de OIOS en agosto de 2007

(2007a) publica un informe titulado "*Valoración de la capacidad y necesidades de evaluación en la Secretaría de las NU*" (*Assessment of evaluation capacities and needs in the United Nations Secretariat*) en el cual indica que todas las agencias que no tengan establecidas políticas de evaluación, las desarrollen en base a las Normas y Estándares de la UNEG<sup>71</sup>.

A pesar de que no existe un consenso generalizado sobre los elementos que deben contener las políticas de evaluación en las instituciones de las NU, la División de Inspección y Evaluación de OIOS en 2007 elaboró, a modo orientativo, unas guías con los elementos generales que conforman la política de evaluación de las agencias pertenecientes al Sistema de las NU (OIOS 2007b). La tabla 4.1 resume los elementos básicos de las políticas de evaluación y su relación con las normas UNEG:

**Tabla 4.1. Políticas de evaluación UNEG**

<b>Elementos básicos de una política de evaluación</b>	<b>Correspondencia con Normas y Estándares UNEG</b>
Explicación sucinta del concepto y rol de la evaluación	Norma 1,3,13 y Estándar 1.2
Estándares generales de evaluación	Norma 11 y Estándares 2.5, 2.6, 2.7, 2.8, 3.10
Clara definición del contexto institucional, roles y responsabilidades	Norma 2, 6, 9 y Estándares 1.1, 1.3, 1.5, 1.6, 1.7, 2.1, 3.10, 3.11
Estrategias de priorización y planificación de las evaluaciones	Norma 7, Estándar 3.1
Organización, gestión y presupuesto de las evaluaciones	Normas 4, 8 y 10, Estándares 3.1 y 3.2
Mecanismos de seguimiento de las evaluaciones	Norma 2, Estándares 1.4, 3.17
Procedimientos para la diseminación	Norma 2, Estándar 1.4

Fuente: Elaboración propia, adaptada de OIOS 2007b

### **Las prácticas evaluativas**

Las prácticas evaluativas deben ser el reflejo de la aplicación de los elementos de la política evaluativa. Así, existen dos prácticas fundamentales que definen el grado de maduración de la función de evaluación en las instituciones de las NU:

- ✓ El sistema de planificación de las evaluaciones (Normas UNEG 2.6 y 4.2) y;

<sup>71</sup> IED-2006-006, recomendación número cinco.

- ✓ El sistema de seguimiento de las recomendaciones de la evaluación (Normas UNEG 12.1, 12.2 y 12.3).

De esta manera, la UNEG especifica que el sistema de planificación de las evaluaciones debe responder a unos criterios específicos previamente identificados que permitan el aprendizaje y la rendición de cuentas. Asimismo, el proceso de planificación ha de ser un proceso consensuado, participativo y, en última instancia, debe ser aprobado al más alto nivel para garantizar el respaldo institucional. Con respecto al sistema de seguimiento, en primer lugar, éste ha de formar parte de la práctica habitual de la institución. Además, la UNEG propone que el ejercicio de dar seguimiento a las recomendaciones emanadas de la actividad evaluativa sea complementado con la provisión de información relativa a dicho proceso de seguimiento a los órganos de gobierno de la organización en cuestión. Finalmente, de acuerdo a la UNEG, se destaca la importancia de alinear las recomendaciones con el ciclo de programación, es decir, considerar las recomendaciones en proyectos futuros donde su aplicabilidad sea pertinente.

Finalmente, como Preskill y Boyle subrayan (2008, pág. 453), *“la existencia de políticas y prácticas evaluativas alineadas con los estándares de calidad evaluativos aumenta las posibilidades de impactos positivos a largo plazo de la práctica evaluativa así como la sostenibilidad de la misma en las organizaciones”*. Esta apreciación es también aplicable al Sistema de las NU.

#### **4.1.3 Cultura de aprendizaje**

Además de los dos elementos descritos hasta ahora, existe un tercer factor explicativo de la capacidad evaluativa que ha ido progresivamente cobrando importancia: la cultura de aprendizaje.

La capacidad de aprendizaje de una organización hace referencia a la capacidad de dicha organización de transformarse y adaptarse a partir de una serie de conocimientos nuevos adquiridos. Una organización que despliega una gran capacidad de aprendizaje es, por tanto, capaz de generar cambios internos que le permitan no sólo adaptarse o anticipar variaciones en el entorno sino también posicionarse mejor en el mismo, agregando valor adicional a su actividad principal de acuerdo a sus potencialidades.

La capacidad de aprendizaje es un concepto complejo al estar éste determinado por una gran cantidad de variables como, por el ejemplo, el entorno externo de la organización (más o menos cambiante), la naturaleza interna de la actividad de la organización, su cultura interna en relación a la aceptación o resistencia al cambio, la estructura organizativa (vertical u horizontal), el liderazgo existente (formal o informal), los incentivos existentes para la innovación, etc.

Todas estas variables van a contribuir a que una organización sea más o menos proclive al aprendizaje y, en consecuencia, a la transformación. Al entender que lo anterior se corresponde con una importante característica de la

capacidad evaluativa, el objetivo del análisis debe ser entonces capturar el grado de capacidad de aprendizaje que despliegan las agencias más importantes del Sistema de las NU y desvelar su relación con la capacidad evaluativa. ¿Hasta qué punto la capacidad de aprendizaje juega un papel relevante en el uso y utilización de las evaluaciones? ¿Es necesario disponer de cierta capacidad de aprendizaje para que las evaluaciones generen resultados que puedan ser utilizados?

Aunque presente en la literatura de evaluación, los conceptos teóricos relativos al aprendizaje organizacional forman parte de la teoría organizativa (Cousins 1996; Huber 1991 y Louis 1994) y están basados en la premisa de que el aprendizaje en las organizaciones es un factor crucial para facilitar procesos de cambios. En este contexto, se entiende que el aprendizaje organizacional no se limita simplemente a la suma del potencial de aprendizaje a nivel individual de las personas que forman dichas organizaciones sino que va más allá.

Herbert Simon (1991, pág. 125) lo resume así:

*“El aprendizaje en el contexto de una organización está muy influenciado por la organización, tiene consecuencias para la organización y va más allá de observar los procesos de aprendizaje en individuos”.*

Robbins (1998) define aprendizaje organizacional como una actividad social donde el conocimiento y las habilidades se ponen en práctica, se critican y son integradas en la organización, constituyéndose como oportunidades para mejorar la efectividad en el ámbito laboral. El aprendizaje supone, así, un aporte activo de las personas las cuales exploran, construyen y descubren nuevos conocimientos desde la práctica diaria.

Como anticipábamos, el paso que distingue conocimiento y aprendizaje es el cambio. El aprendizaje organizacional implica que las organizaciones son capaces de modificar sus estructuras y sus programas de acción sobre la base de nuevos conocimientos adquiridos.

Aunque no forma parte del foco central de análisis en la literatura sobre evaluación, la conexión entre el aprendizaje organizacional y la práctica evaluativa ha sido subrayada por algunas autoras y autores entre los que se encuentran Preskill y Torres. Estas evaluadoras establecen que un continuo proceso de crecimiento y mejora en cualquier organización se sustenta a partir de la información nueva generada por la organización para retroalimentar procesos y resultados como, por ejemplo, hallazgos de la evaluación, que originan cambios dentro de la misma. Además, establecen que la posibilidad de cambio y transformación está integrada en las actividades del trabajo y en la propia estructura organizacional (por ejemplo, en su cultura, sistemas, formas de liderazgo, y mecanismos de comunicación), abogando, por consiguiente, por un lineamiento de valores, actitudes y percepciones entre sus miembros (Preskill y Torres 1999, Torres et al 1996).

De acuerdo a las Normas y Estándares UNEG, algunos de los indicadores que reflejan una mayor cultura de aprendizaje en el ámbito de la evaluación son:

- ✓ Diseminación de la información evaluativa
- ✓ Participación en la evaluación
- ✓ Conocimiento evaluativo

Más específicamente, en las normas UNEG N2.7, N10.1 y N10.2, se enfatiza la necesidad de establecer mecanismos para la diseminación y acceso al conocimiento de información evaluativa, a la vez que transparencia, consulta y participación de los actores clave en todas las fases de la evaluación.

En base a lo anterior, se puede concluir que la capacidad de aprendizaje de una organización entendida como la capacidad de dicha organización de adaptarse, cambiar y/o transformarse de acuerdo a nuevos flujos de información, es un criterio que, aunque no siempre de forma explícita, se encuentra presente y subyace en el marco teórico de capacidad evaluativa; y su importancia será analizada posteriormente en el apartado 5.1 y 5.2.

#### **4.2. La hipótesis y su marco de análisis en el Sistema de las NU**

El concepto de capacidad evaluativa en el Sistema de las NU ha cobrado una gran relevancia en los últimos años, especialmente tras los compromisos adquiridos entre los países donantes y los países en desarrollo reflejados en la Declaración de París y el Programa de Acción de Accra ya mencionados (apartado 2.2). En dichos eventos organizados por la OCDE se subraya la necesidad de aunar esfuerzos para mejorar la efectividad de la ayuda y, en ese contexto, se señala a la capacidad evaluativa como un instrumento de gran relevancia para que los países socios se comprometan con los procesos de planificación, monitoreo y evaluación relativos a las contribuciones realizadas en el marco de la cooperación internacional.

La función de evaluación en las NU se ha visto así fortalecida sustancialmente, sobre la base de principios fundamentales como la independencia, credibilidad y utilidad de la función de evaluación suponiendo un mayor desarrollo de la capacidad evaluativa. En cualquier caso, a pesar de los esfuerzos por fortalecer la capacidad evaluativa llevados a cabo hasta la fecha, los informes de seguimiento publicados por organismos de control interno en el marco de las NU coinciden en que la capacidad evaluativa existente es aún deficiente<sup>72</sup>.

En base a la teoría previamente descrita, podemos intuir que los tres factores identificados (institucionalización, políticas/prácticas y capacidad de aprendizaje) deberían contribuir al logro de los objetivos inherentes a la

---

<sup>72</sup> Como ya se ha indicado, OIOS realiza el análisis de la función de evaluación de los organismos que reportan a la Secretaría de manera bianual. El último fue publicado en marzo de 2015 (A70/72). Además, JIU también ha realizado en 2014 una valoración sobre las funciones de evaluación de los órganos supervisados bajo su mandato. Ambos informes son públicos y accesibles en sus respectivas páginas web:

<https://oios.un.org/page?slug=evaluation-report>

[https://www.unjiu.org/en/reports-notes/JIU%20Products/JIU\\_REP\\_2014\\_6\\_English.pdf](https://www.unjiu.org/en/reports-notes/JIU%20Products/JIU_REP_2014_6_English.pdf)

capacidad evaluativa (calidad y uso). Aun siendo posible que existan otros factores, al menos éstos tres parecen ser fundamentales de acuerdo con los modelos de referencia existentes en la actualidad y analizados previamente.

Así, el Sistema de las NU inspirado por el marco teórico ha dirigido importantes esfuerzos para establecer funciones de evaluación independientes, dotar de recursos específicos a dichas unidades y establecer políticas y prácticas que faciliten la operatividad de la evaluación en las organizaciones del sistema.

Sin embargo, existen algunos interrogantes sin contestar que han motivado esta investigación: ¿Porqué tras tantos esfuerzos orientados a reforzar la capacidad evaluativa como mecanismo de cambio y mejora en las NU, los resultados obtenidos no son satisfactorios? ¿Cómo se explica la percepción generalizada de que la práctica evaluativa en las NU aún no ha conseguido influenciar de manera significativa el proceso de toma de decisiones estratégicas dentro de sus organismos?

Analizando los interrogantes anteriores, el grado de importancia que cada factor tiene de cara a mejorar la capacidad evaluativa emerge como un área pendiente de investigar y en la cual se pueden encontrar respuestas. La razón fundamental de lo anterior reside en que el marco teórico existente no pondera la importancia relativa de los factores que identifica como catalizadores de capacidad evaluativa. Hasta la fecha, el marco teórico se concentra fundamentalmente en apuntar hacia los factores donde debemos realizar el análisis y dirigir los esfuerzos para lograr un grado de capacidad evaluativa adecuada que optimice la función de evaluación en una organización, pero no analiza la importancia de cada factor, bien de forma individual o en su relación con los demás factores.

La hipótesis que se contrasta en esta investigación se centra en la idea de que la cultura de aprendizaje organizacional es un elemento clave dentro de las instituciones para favorecer la utilización de la evaluación. Sin embargo, la cultura de aprendizaje organizacional por si sola puede no ser suficiente para realizar una práctica evaluativa de calidad y sostenible en las instituciones. Esta cultura de aprendizaje debe ir acompañada de un sólido marco normativo en materia de evaluación, para convertirse en un factor fundamental para el logro de la capacidad evaluativa que se necesita en las instituciones de las NU, es decir, una capacidad evaluativa que permita la mejora de la calidad de las evaluaciones y su influencia en la toma de decisiones. Hasta ahora, esta combinación de factores que resulta de la cultura de aprendizaje organizacional junto con el marco normativo, no ha sido reconocida por los modelos teóricos que analizan la capacidad evaluativa como un elemento de crucial importancia.

Así, la hipótesis central a contrastar es la siguiente: la capacidad de aprendizaje se constituye como condición necesaria para lograr resultados satisfactorios de utilización de la evaluación, y si ésta es combinada con un marco normativo adecuado, constituye entonces una fórmula efectiva para producir evaluaciones de calidad en las organizaciones.

Sin cultura de aprendizaje, el resto de factores difícilmente van a lograr por sí mismos la plena utilización de la función de evaluación. Así, deficiencias en la cultura de aprendizaje producen efectos negativos en el logro de los resultados que persigue la capacidad evaluativa, a pesar de que las otras condiciones del marco teórico estén presentes (estructuras, recursos y políticas/prácticas evaluativas). La cultura de aprendizaje es, por tanto, un eje central en la capacidad evaluativa y su presencia permite que los demás factores facilitadores sean efectivos.

En esta investigación, el análisis y contraste de estas hipótesis se realiza sobre el Sistema de las NU, dado que este Sistema representa un amplio marco de acción internacional y nacional en el que la evaluación juega un papel fundamental. Los proyectos y programas, que, como se ha mencionado anteriormente (apartado 3.1), se constituyen en una de las formas de intervención más representativas de los organismos del Sistema de las NU, incorporan en el ciclo de proyecto la propia práctica evaluativa como un elemento intrínseco de los mismos, factor que junto a otros muchos también ya mencionados en el apartado 3.2 (como, por ejemplo, la existencia de OIOS o la JIU) reflejan la importancia que se le otorga a la evaluación dentro de la práctica rutinaria de las NU. El Sistema de las NU constituye así un importante y representativo caldo de cultivo para el análisis de la práctica evaluativa y, por ende, de la capacidad evaluativa. Hallazgos en este terreno, por tanto, pueden contribuir a mejorar la comprensión existente sobre la capacidad evaluativa no solo en el Sistema de las NU que va a ser analizado, sino también en otro tipo de organizaciones donde la función de evaluación sea de gran relevancia, y en las que, en consecuencia, se destinen esfuerzos al impulso de su correspondiente capacidad evaluativa.

#### **4.3. Metodología propuesta**

Para poder contrastar la hipótesis planteada se va a realizar un análisis comparativo entre diez organismos que componen una muestra representativa del Sistema de las NU. Este análisis, centrado en la capacidad evaluativa de las agencias, tratará de poner de manifiesto que, efectivamente, aquellos organismos que gozan de mayor capacidad de aprendizaje y, además, han establecido políticas y prácticas evaluativas de manera consistente, realizan evaluaciones de mayor calidad que proporcionan resultados y hallazgos útiles para la gerencia permitiendo a la organización en cuestión llevar a cabo cambios estructurales y/u operativos con mayor facilidad.

Las 10 entidades de las NU que forman parte de la muestra seleccionada gozan todas ellas de una importante tradición en relación al desarrollo e implementación de la práctica evaluativa. Estos organismos han realizado procesos de revisión por pares (*peer reviews*) de sus funciones de evaluación que permiten realizar un análisis exhaustivo del estado de dichas instituciones en relación a los factores desencadenantes de capacidad evaluativa y a los resultados obtenidos en este terreno (calidad y uso).

Pero, además para seleccionar la muestra se ha utilizado un segundo criterio de selección: se ha optado a propósito por incluir una representación de los

diferentes organismos, tal y como han sido expuestos en el apartado 3.2, con el fin de abarcar la evaluación en diferentes tipos de organizaciones dentro de NU, y con ello, reflejar así su heterogeneidad. La muestra está compuesta por:

- ✓ Organismos donde la evaluación es parte intrínseca de su mandato (OIOS)
- ✓ Organismos que reportan a la Secretaría (ONU Mujeres, UNO Habitat)
- ✓ Programas y Fondos (PNUD, UNICEF, PMA, PNUMA)
- ✓ Organismos/entidades especializadas (ONUDI, FMAM<sup>73</sup>, FAO)

El método seleccionado para realizar la comparación se denomina Análisis Comparativo Cualitativo (*Qualitative Comparative Analysis*, QCA). Esta metodología presenta una serie de características que la hacen especialmente adecuada e interesante para el análisis propuesto en esta investigación.

Las raíces de este método de investigación se encuentran en el 'método de la concordancia' (*method of agreement*) y el 'método de la diferencia' (*method of difference*) de los cánones de J.S Mills (1967<sup>74</sup>). Según el método de la concordancia, si ante el análisis de situaciones comparadas, nos encontramos con que éstas solo tienen una circunstancia en común, se puede afirmar que esta circunstancia común es la causa o el efecto del fenómeno bajo análisis. Por contraste, el método de la diferencia esgrime que si cuando comparamos ciertas situaciones nos encontramos con que todas sus variables coinciden excepto una, ésta se convierte en la causa o el efecto del fenómeno en cuestión. Ambos métodos buscan el establecimiento de relaciones entre ciertos patrones de conducta y tratan de buscar elementos de relación causal, eliminando el resto de variables irrelevantes.

El Análisis Cualitativo Comparativo (QCA, por sus siglas en inglés) es un método de investigación desarrollado por Charles Ragin en 1987 que ofrece una alternativa sólida a la disyuntiva que se presenta frecuentemente en las ciencias sociales de la necesidad de tener que elegir entre métodos cuantitativos y métodos cualitativos de investigación. De esta forma, QCA utiliza un riguroso método cuantitativo para el análisis comparativo, en el cual intervienen, a su vez, elementos cualitativos al habilitar durante el proceso comparativo la consideración del conocimiento cualitativo del investigador/a sobre la materia en cuestión.

QCA es, además, un método que ha contribuido, de manera esencial, al desarrollo de la metodología de las ciencias sociales puesto que ofrece en el ámbito del análisis comparativo, un método sistemático y riguroso (Schneider y Wagemann 2012).

QCA se basa en el análisis de relaciones de fundamentación teórica (*theory-driven*) y se centra principalmente en establecer conexiones entre condiciones causales y resultados específicos. La lógica de las conexiones entre condiciones y resultados se establece a través de patrones de necesidad o suficiencia de las condiciones, tanto a nivel individual, como de una

---

<sup>73</sup> A través de su vinculación al Grupo del Banco Mundial.

<sup>74</sup> El original fue publicado en 1843.

combinación de las mismas<sup>75</sup>. Es, por ello, que, a diferencia de otros métodos, QCA parte de la premisa de la complejidad causal a la hora de realizar el análisis comparativo (Mahoney y Goertz 2006).

En esta investigación, las comparaciones se realizan entre entidades o casos (organismos del Sistema de las NU) a partir de la transformación de los mismos en diferentes configuraciones o grupos de variables, componentes o - por ser fieles a la terminología empleada en QCA – **condiciones**, que, agrupadas de una u otra manera, contribuyen a generar un resultado determinado de interés para la investigación en cuestión. En este análisis, las condiciones vienen representadas por los factores que componen la capacidad evaluativa y los resultados que suscitan interés son la calidad y el uso de la evaluación. Cada organismo de la muestra despliega una calificación determinada de pertenencia en relación tanto a las condiciones como a los resultados. De esta forma, podremos comparar, por ejemplo, organismos que muestren un resultado positivo (una gran calidad de las evaluaciones y un uso elevado de las mismas) a partir de una combinación determinada de condiciones, con otros organismos cuyo resultado no sea tan alentador (sus evaluaciones no tienen tanta calidad o no influyen tanto en la organización) y que despliegan una combinación diferente a la anterior en cuanto a las condiciones presentes.

Más específicamente, en el presente trabajo se han identificado 3 condiciones que emanan del marco teórico como factores explicativos de la capacidad evaluativa, concretamente:

- ✓ Institucionalización de la función de evaluación
- ✓ Presencia de políticas y prácticas evaluativas
- ✓ Capacidad de aprendizaje de la entidad

A su vez se estudiará en cada agencia, el efecto de estas condiciones (de forma individual y combinada) sobre la consecución de los dos resultados esperados que se desgranar de la definición teórica comúnmente aceptada para la capacidad evaluativa, y que son:

- ✓ Calidad de las evaluaciones
- ✓ Uso de la evaluación

La fuente de información utilizada para cada uno de los diez organismos objeto de estudio son los informes de revisión por pares (*peer reviews*) realizados sobre sus funciones de evaluación. Estos informes, descritos en el apartado 3.4.2, son realizados por paneles compuestos de evaluadores/as independientes y provenientes de otras agencias de NU y otros organismos vinculados a la cooperación bilateral o multilateral. El carácter independiente de los informes y su proceso sistemático de elaboración (similar al de una evaluación), que incluye múltiples instrumentos de recogida y contrastación de

---

<sup>75</sup>A efectos terminológicos, hemos denominado condición a la configuración de condiciones bien sea a nivel individual como combinada.

datos (entrevistas, encuestas, visitas de campo, análisis documental), hacen de estos informes unas herramientas objetivas, creíbles y, en definitiva, de gran valor para nuestro análisis.

A través del método QCA podremos comparar los diferentes casos (entidades), a partir sus indicadores (obtenidos del análisis teórico de los factores explicativos de la capacidad evaluativa), desvelando las combinaciones de condiciones que son necesarias y/o suficientes para la obtención del resultado de interés (calidad y uso de las evaluaciones).

Una condición será considerada **necesaria** si ésta se encuentra siempre presente cuando ocurre el resultado. Es posible que también se encuentre presente en circunstancias en las que no se produzca el resultado aunque, si éste se produce, la condición se encuentra presente.

En contraste, una condición es considerada como **suficiente** cuando su presencia implica que el resultado se produce, aunque también puede ocurrir que, produciéndose el resultado, la condición no esté presente ya que éste se logra por otros medios.

Este tipo de análisis nos permite entender hasta qué punto una condición juega un rol determinante en la consecución de un resultado particular. En nuestra investigación, el marco teórico existente no ha ponderado hasta qué punto los factores (que conforman la capacidad evaluativa bien de forma individual o en combinación con otros factores) son necesarios y/o suficientes para que las evaluaciones tengan calidad y sean útiles e influyentes en la toma de decisiones de las organizaciones.

La complejidad causal del QCA se articula en base a tres principios (Schneider y Wagemann 2012):

- ✓ Equifinalidad: Diferentes 'camino's o condiciones causales pueden llevar al mismo resultado. En este sentido, puede existir más de una condición suficiente para producir el resultado.
- ✓ Causalidad coyuntural: Las condiciones pueden ser analizadas desde un punto de vista individual o desde una perspectiva combinada de condiciones. De hecho, a veces, una condición por sí sola no es suficiente para influenciar en el resultado, pero sí lo es, en cambio, si ésta se presenta en combinación con otras condiciones.
- ✓ Asimetría: el hecho de que una condición origine un resultado determinado no implica necesariamente que, ante la ausencia de dicha condición, no se produzca el resultado. Así, el conocimiento de las causas que producen un resultado no implica necesariamente que conozcamos las razones por las cuales el resultado no se produce.

El QCA es considerado una familia de métodos. La primera versión de QCA es el denominado hoy en día *crisp-set* QCA (csQCA), método que aparece en 1987, aunque fue nombrado de esta manera cuando más adelante se introdujeron nuevas variantes del QCA.

El csQCA establece parámetros de dicotomía, tanto de las condiciones como del resultado. Este hecho ha provocado que csQCA recibiera muchas críticas debido a la pérdida innecesaria de información. En general, se puede afirmar que existen situaciones donde es posible dividir la realidad en dos posibles opciones pero, en muchas ocasiones, la dicotomía puede llevar a situaciones algo artificiales, y difícilmente aplicables a la realidad. Para poder incluir la posibilidad de varias opciones dentro de una misma variable, Ragin en el año 2000 desarrolló el fsQCA que es una combinación de teoría de lógica difusa (*fuzzy sets*) y QCA. La principal diferencia estriba en que, en la lógica difusa, se permite una pertenencia parcial a un conjunto. Con fsQCA no es necesaria la dicotomía en el concepto a analizar. Se puede afirmar que fsQCA ha incrementado la popularidad y aplicación del QCA como herramienta comparativa cualitativa frente a csQCA (Pennings 2003; Rihoux y Grimm 2006; Schneider y Wagemann 2007; Sager 2005)<sup>76</sup>.

En fsQCA, para expresar el grado de pertenencia a un conjunto, se construyen las *fuzzy scales*, o escalas difusas que atribuyen a los casos valores entre 0 y 1 en la lógica Booleana y difusa (Ragin 2000). La pertenencia de cada caso en el conjunto de las condiciones puede variar entre pertenencia total (valor 1), y no pertenencia (valor 0). En particular:

- ✓ Valor 1 es el umbral para considerar plena inclusión en un conjunto.
- ✓ Valor 0 es el umbral para considerar plena exclusión de un conjunto.
- ✓ Valor 0.5 es el punto de intersección de máxima ambigüedad, o lo que es lo mismo, aquel valor para el cual el grado de indeterminación respecto a la potencial inclusión del caso en un conjunto resulta máximo (Ragin 2008, Verkuilen 2005).

El proceso de construcción de escalas difusas y su posterior ‘traspolación’ de valores difusos se denomina calibración. Ragin específicamente señala la importancia de presentar la calibración de manera transparente, de tal manera que los datos y la metodología puedan ser contrastados, mejorados, y replicados en el futuro (Ragin y Sonnett 2005). La calibración se realiza a partir de las asunciones básicas sobre la pertenencia a un conjunto (*set membership*). Estas asunciones deben estar basadas en el marco teórico, el conocimiento de los casos y la información empírica existente.

Es importante resaltar que, si bien los valores de pertenencia a un conjunto se encuentran siempre entre 0 y 1, éstos no representan probabilidades sino su grado de integración en el conjunto. El valor *fuzzy set* de una determinada

---

<sup>76</sup> De hecho, y puesto que con csQCA está limitado únicamente a dos valores, 0 y 1, se puede afirmar que csQCA es un caso especial de fsQCA.

condición es representado a través de un valor que mide su pertenencia a un conjunto, no una probabilidad (Cioffi-Revilla 1981).

Como se ha anticipado, las condiciones propuestas y los resultados, en nuestra investigación, tienen su justificación en la definición de capacidad evaluativa y en el marco teórico explicado anteriormente. En este sentido, el factor contextual cobra especial relevancia en la calibración de las condiciones (Collier 1998, Ragin 2000). Dado que la tesis se enfoca en el contexto del Sistema de las NU, para cada una de las condiciones y resultados propuestos, se ha procedido a un análisis profundo de su aplicación dentro de las Normas y Estándares UNEG.

Además, puesto que nuestro objeto de estudio se basa en datos cualitativos, el método de calibración utilizado es el denominado método indirecto. En contraposición con el método directo de calibración que asigna un puntaje en escala difusa a partir de valores numéricos, el método indirecto se basa en la identificación de un grupo de características cualitativas que definen o explican la condición; y de una escala que indica el grado de cumplimiento de dichas características cualitativas.

Conceptualmente, el método indirecto de calibración presenta similitudes con otros métodos cualitativos de análisis, como el desarrollado por Leithwood y Montgomery en 1987 llamado perfil de innovación (*innovation profile approach*), y posteriormente utilizado por Bourgeois y Cousins (2013). Este enfoque, aplicado en el sector educativo, se basa en identificar cambios observables dentro de una misma variable y establecer una escala gradual. Para ello, se definen ciertos indicadores cualitativos para las variables, que deben ser graduados en torno a una escala que debe cubrir tanto la situación actual como la óptima o deseada. La principal dificultad estriba en la identificación de los indicadores y de su escala puesto que el paso de un estado a otro dentro de la escala debe ser justificado en base a cambios observables y verificables. Para nuestro análisis sobre la capacidad evaluativa, cada una de las tres condiciones y de los dos resultados se ha desglosado en características o factores que emanan del marco teórico y las Normas y Estándares UNEG.

Se ha definido así una escala difusa de cuatro valores, tal y como se define a continuación (Ragin 2009):

- 0.0 - Completamente fuera del conjunto
- 0.33 - Mayoritariamente fuera del conjunto
- 0.66 - Mayoritariamente dentro del conjunto
- 1.00 - Completamente dentro del conjunto

En base a la escala anterior se ha elaborado la Tabla 4.2 donde se presentan las condiciones y los resultados para el análisis, así como las características cualitativas (selección de indicadores) que comprende cada condición y cada resultado, utilizando como marco de referencia para la construcción de los indicadores todo lo desarrollado en el marco teórico y en la revisión de la literatura sobre capacidad evaluativa y su aplicación a las Normas UNEG.

Tabla 4.2. Indicadores y su graduación en escala difusa

CONDICIONES	Factor clave	0.00	0.33	0.66	1.00
Cultura de aprendizaje (CA)	<p>Existencia y grado de aplicación de una estrategia de disseminación de la información (incluyendo diversas actividades como talleres de disseminación, publicaciones dirigidas a audiencias seleccionadas, traducciones, acceso público de informes etc.)</p>	<p>Inexistencia de una estrategia de disseminación de información evaluativa. Informes no accesibles públicamente, inexistencia de información sobre las actividades de la unidad de evaluación.</p>	<p>Existencia de ciertos signos de disseminación de la información, pero principalmente aislados y <i>ad hoc</i> (documentos públicos, talleres en solo algunas evaluaciones), no responden a una estrategia consistente y aplicable para todas las evaluaciones.</p>	<p>Clara evidencia de aplicación de una estrategia de disseminación, aunque se han identificado algunas limitaciones que hacen que su aplicación no sea plena (por ejemplo una limitada comunicación con oficinas de campo, documentos no traducidos)</p>	<p>Plena existencia y aplicación de una estrategia de disseminación de información evaluativa (acceso público, productos diseñados para diferentes audiencias, traducción, realización consistente de talleres divulgativos con amplia presencia de actores clave, etc.)</p>
	<p>Grado de participación y consulta del personal clave en la evaluación y en sus diferentes fases (TdRs, informe, metodología, etc.)</p>	<p>El personal del programa evaluado no es consultado en el proceso de evaluación, no tienen oportunidades en general de aportar sugerencias a los TdRs ni a los informes de evaluación.</p>	<p>Existen ciertos signos de consulta y participación, pero generalmente el nivel de participación es inconsistente (no hay momentos clave del proceso donde se produzca la consulta, ni procedimientos para la identificación de actores clave). Los propósitos de consulta son principalmente informativos con poca capacidad de incidencia.</p>	<p>El nivel de consulta y participación es alto, se aplica consistentemente a ciertas fases clave del proceso (metodología, TdRs, informes). En general, los comentarios son tomados en cuenta. Aun así, se observan ciertas limitaciones, principalmente ligadas a la escasa cobertura del proceso consultivo y limitada capacidad de incidencia.</p>	<p>El nivel de consulta y participación es muy alto, consistente, y coherentemente aplicado en todas las fases. Existe un claro criterio establecido en la identificación de actores clave, y los comentarios externos son altamente considerados en la toma de decisiones ligada a la evaluación.</p>
	<p>Nivel de conocimiento y sintonía con los propósitos de la evaluación y sus potenciales usos con respecto al aprendizaje organizacional.</p>	<p>El personal desconoce y confunde cuál es el propósito de la evaluación y qué puede aportar de valor añadido al rendimiento organizacional.</p>	<p>El personal tiene cierto conocimiento sobre las actividades relativas a la evaluación, pero con frecuencia son confundidas con otras tareas relacionadas, por ejemplo, monitoreo y auditoría. El factor aprendizaje tiende a quedar en un papel muy poco relevante.</p>	<p>A nivel general, se observa alto conocimiento sobre lo que la función de evaluación aporta. Percepción positiva sobre el valor añadido de la evaluación, el concepto de aprendizaje prima sobre la rendición de cuentas. Se observa cierto desnivel de conocimiento con respecto a personal de las oficinas de campo y otros actores clave.</p>	<p>Conocimiento alto y consistente de la evaluación entre el personal y confirmado también por otros actores clave (órganos de gobierno). La evaluación es percibida muy positivamente como elemento que favorece el aprendizaje y la mejora organizacional. Tiene un papel relevante en la organización.</p>
<p>CA1 Diseminación de la información evaluativa (N2.7, N13.1, N13.2)</p>					
<p>CA2 Participación en la evaluación (N5.1, N10.1, N10.2)</p>					
<p>CA3 Motivación/cultura evaluativa (N1.1, N1.5)</p>					

CONDICIONES	Factor clave	0.00	0.33	0.66	1.00
Institucionalización de la función de evaluación (INS)	<p>Grado de independencia de la unidad de evaluación con respecto al resto unidades programáticas. Grado de autoridad y autonomía para someter información evaluativa directamente a los órganos de gobierno.</p> <p><b>INS1</b> Estructura (N2.2, N5.2, N6.1)</p>	<p>La unidad no es independiente de las unidades programáticas. No existe canal de comunicación directo entre la unidad y los órganos de gobierno.</p>	<p>La unidad es estructuralmente independiente de las unidades programáticas. Sin embargo, no tiene canal directo de comunicación a los órganos de gobierno.</p>	<p>La unidad es estructuralmente independiente de las unidades programáticas. Tiene canal directo de comunicación a los órganos de gobiernos. Sin embargo, a nivel jerárquico, la jefatura de la unidad de evaluación reporta a la dirección de la organización, no a los órganos de gobierno.</p>	<p>La unidad es estructuralmente independiente de las unidades programáticas. Tiene canal directo de comunicación con los órganos de gobiernos. Además, la jefatura de la unidad de evaluación reporta directamente a los órganos de gobierno, quienes la eligen por un periodo de tiempo determinado.</p>
	<p>Grado de transparencia en la asignación de presupuesto para evaluación, nivel de suficiencia para cubrir el plan de evaluación, y adecuación del presupuesto para cubrir las necesidades de la unidad de evaluación y planificar a largo plazo.</p> <p><b>INS2</b> Recursos financieros (N2.3, N4.1)</p>	<p>La asignación de presupuesto no sigue criterios transparentes y los datos son difícilmente accesibles. Se considera que el presupuesto no es suficiente para cubrir el plan de evaluación, ni las necesidades operativas de la unidad de evaluación.</p>	<p>Existen datos disponibles sobre el presupuesto pero las partidas asignadas a evaluación no son claras (mezcladas con otras partidas, como monitoreo o auditoría). El plan de evaluación excede el presupuesto regular, y, por tanto, gran parte del plan depende de partidas extrapresupuestarias no predecibles.</p>	<p>Existen datos disponibles y un criterio transparente de asignación. El plan de evaluación es factible puesto que puede ser cubierto con recursos predecibles y, en general, el presupuesto es adecuado para cubrir las necesidades de la unidad. Los planes son aprobados bi-anales o anuales, lo que no facilita la planificación a nivel estratégico.</p>	<p>Alto nivel de transparencia en la asignación de recursos, datos fácilmente accesibles y claros. El presupuesto permite cubrir el plan de evaluación y hacer frente a las necesidades de la unidad de evaluación con total autonomía e independencia. Es posible realizar planificación estratégica, puesto que el presupuesto es aprobado por periodos multianuales (3-4 años).</p>
	<p>Balance adecuado de posiciones sénior-junior, <i>staff</i>-consultores, y grado de suficiencia de los recursos humanos para hacer frente al plan de evaluación.</p> <p><b>INS3</b> Recursos humanos (N2.4, N2.5, N6.3, N9.1, N9.2, N9.3)</p>	<p>Personal insuficiente para hacer frente al plan de evaluación. No hay balance entre sénior y junior <i>staff</i>. El número de personas externas es mucho más elevado que el personal propio por limitaciones presupuestarias, y la capacidad de gestionar los proyectos en la oficina es deficiente.</p>	<p>Existe balance entre senior y junior <i>staff</i>, pero se tiende en exceso a la contratación de personas externas para poder hacer frente al plan de evaluación, y, en ocasiones, se les otorga labores gerenciales. En agencias donde la evaluación es descentralizada, se observa gran desequilibrio entre las capacidades a nivel de HQs y oficinas de campo.</p>	<p>Existe balance entre sénior y junior <i>staff</i>, y entre personal interno y externo. Las personas externas se encargan del apoyo técnico, no de labores gerenciales. En agencias de evaluación descentralizada, hay equilibrio entre las capacidades de la oficina central y las oficinas de campo. Aun así, se observan dificultades para poder llevar a cabo el plan de evaluación, limitada disponibilidad de recursos humanos o restricciones en la contratación de consultores/as.</p>	<p>Existe balance adecuado de posiciones sénior y junior, <i>staff</i> y no <i>staff</i>. Las personas externas tienen claramente asignadas labores de apoyo técnico. En agencias donde la evaluación es descentralizada existen capacidades muy equilibradas entre la oficina central y oficinas de campo. Los recursos humanos son considerados adecuados y suficientes para poder llevar a cabo el plan de evaluación.</p>

CONDICIONES	Factor clave	0.00	0.33	0.66	1.00	
<b>Políticas y prácticas evaluativas (PP)</b>	<b>PP1</b> Plan de evaluación (N2.6, N4.1, 4.2)	Existencia de criterio multidimensional y transparente en la selección de temas, nivel de consulta y participación de actores clave en la elaboración del plan, y grado de autonomía de la unidad de evaluación para discutir y aprobar directamente con los órganos de gobierno.	No existe un plan de evaluación, ni se conocen criterios predeterminados para proceder con una evaluación. El grado de consulta entre actores clave es mínimo, y no existen canales consistentes de aprobación de los temas a evaluar.	Puede existir cierto criterio en la selección de temas, pero no es un criterio aplicado consistentemente, ni tiene carácter multidimensional. El proceso de consulta tiende a ser limitado, y no se aplica de manera consistente para todas las evaluaciones. El plan es aprobado a nivel interno (ED), no necesariamente por los órganos de gobierno.	Existe un criterio transparente y multidimensional en la selección de temas a incluir en el plan de evaluación. El proceso de consulta tiende a incluir a los principales actores clave de manera consistente. Se pueden observar ciertas inconsistencias, o bien en el proceso de consulta, o en el proceso de aprobación del plan, a veces limitado a nivel interno (no órganos de gobierno)	Existe un criterio transparente y multidimensional en la selección de temas a incluir en el plan de evaluación, y el proceso de consulta incluye a los principales actores clave de manera consistente. El proceso de aprobación del plan incluye discusiones con los órganos de gobierno, quienes aprueban el plan en última instancia.
	<b>PP2</b> Política de evaluación (N3.1)	Existencia de política de evaluación que se adhiere a las normas UNEG y canales de aprobación.	No existe política de evaluación, ni documentos que puedan suplir la	No existe política en si, pero si ciertos documentos (guías, procedimientos), que hacen de política. Se observan ciertas inconsistencias con las normas UNEG y su aprobación es interna.	Existe política de evaluación, o al menos documentos que implícitamente hacen de política, y , en general, se adhiere/n a las normas UNEG. Se pueden observar ciertas limitaciones en su proceso de aprobación, solo a nivel interno (no órganos de gobierno)	Existe política de evaluación y, en general, se adhiere a las normas UNEG. La política es aprobada en última instancia por los órganos de gobierno.
	<b>PP3</b> Seguimiento de las recomendaciones (N2.6, N12.1, N12.2, N12.3)	Existencia de un sistema de seguimiento de las recomendaciones, grado de reporte a los órganos de gobierno, grado de utilización en la formulación del nuevo ciclo programático.	No existe un sistema de seguimiento a las recomendaciones, ni análisis que pueda ser preparado para discusión de los órganos de gobierno. No hay constancia de que las recomendaciones se usen en la formulación del nuevo ciclo.	Existe un sistema incipiente de seguimiento a las recomendaciones, aunque es inconsistente (no se aplica a todas las evaluaciones). La información que se analiza y se lleva a los órganos de gobierno es parcial y limitada, y no hay evidencia de que se use en la formulación del nuevo ciclo.	Existe un sistema de seguimiento a las recomendaciones y se aplica a todas las evaluaciones. La información se analiza y discute a nivel interno. No se lleva necesariamente a los órganos de gobierno. Podrían mejorarse ciertos aspectos del sistema (más importancia a aspectos cualitativos). Existen signos de que se usa en la formulación del nuevo ciclo.	Existe un sistema de seguimiento a las recomendaciones y se aplica a todas las evaluaciones. La información se analiza y discute a nivel interno y externo (órganos de gobiernos). Existen claras evidencias de uso en la formulación del nuevo ciclo.

RESULTADOS	Factor clave	0.00	0.33	0.66	1.00	
<b>Calidad (QUA)</b>	<b>QUA1</b> Informes y estándares (N8.2)	Grado de calidad de los informes en función de los estándares de calidad establecidos por UNEG.	Calidad muy deficiente de los informes de evaluación en base a estándares UNEG	Calidad inconsistente. Se encuentran importantes carencias metodológicas, falta de robustez de las conclusiones, factibilidad de las recomendaciones.	La calidad en general es aceptable y en base a normas UNEG. Pueden encontrarse ciertas limitaciones, no graves, de los informes a mejorar.	La calidad es muy aceptable, y en base a normas UNEG. Destaca la gran solidez y rigor metodológico.
	<b>QUA2</b> Control de calidad (N8.1)	Existencia y aplicabilidad de controles internos que garantizan que la evaluación se lleva a cabo en base a procedimientos de calidad.	Controles de calidad inexistentes.	Existen ciertos controles de calidad, pero son considerados 'light', y no se aplican consistentemente.	Existen controles de calidad en base a procedimientos rigurosos. Se observan ciertas inconsistencias en cuanto a su aplicabilidad (en particular evaluación descentralizada)	Existen controles de calidad en base a procedimientos rigurosos. Se aplican consistentemente en todas las evaluaciones.
	<b>QUA3</b> Valoración positiva de los actores clave	Nivel de satisfacción de los actores clave sobre la calidad de los informes de evaluación	Percepción muy negativa. Los informes son considerados poco creíbles, con importantes deficiencias.	Percepción muy desigual sobre la calidad de los informes, particularmente negativa del personal interno.	Percepción generalizada bastante positiva de la calidad de los informes, aunque a veces a nivel interno pueden existir algunas críticas.	Percepción generalizada muy positiva de la calidad de los informes tanto a nivel externo como interno.
<b>Uso (USO)</b>	<b>USO1</b> Generación de cambios (N4.1)	Existencia de casos reportados de cambios generados como consecuencia de una evaluación	No hay casos reportados y, en general, una percepción negativa generalizada sobre el uso de la evaluación como instrumento de cambio.	Existe la percepción de que la evaluación puede generar cambios e incluso algunos cambios detectados a nivel operativo, pero se observa cierta receptividad del personal interno a adoptarlos.	Existen casos de cambios generados como consecuencia de la información evaluativa, y actitud positiva del <i>staff</i> a adoptarlos, pero tienden a ser cambios operativos (de procedimiento), no a nivel de estrategias o políticas.	Existen casos de cambios generados como consecuencia de la información evaluativa, tanto a nivel estratégico como operativo, y alta receptividad del <i>staff</i> a adoptarlos.
	<b>USO2</b> Implementación de recomendaciones (N2.6, N12.1, N12.2, N12.3)	Porcentaje de recomendaciones implementadas y grado de empoderamiento del personal interno en su implementación	Porcentaje muy deficiente de recomendaciones implementadas, inexistencia de sistema de seguimiento.	Porcentaje insuficiente, sistema incipiente y parcial de las recomendaciones. No empoderamiento del <i>staff</i> .	Porcentaje suficiente y aceptable de las recomendaciones, todavía con aspectos a mejorar, principalmente la implementación de recomendaciones de tipo estratégico.	Porcentaje altamente satisfactorio de las recomendaciones, tanto de recomendaciones de tipo estratégico como operativo.
	<b>USO3</b> Presencia en la toma de decisiones (N2.7)	Presencia y relevancia de la evaluación en procesos y foros de decisión.	La evaluación y sus resultados no se discuten a nivel de gerencia y su presencia es muy poco relevante en foros de decisión.	La presencia de la evaluación en foros de decisión es limitada y su papel poco relevante.	Se considera aceptable el nivel de presencia y relevancia de las evaluaciones en los foros de decisión, aunque se observa cierta tendencia a los procesos internos, no tanto con órganos de gobierno.	Se considera altamente satisfactorio el nivel de presencia y relevancia de las evaluaciones en los foros de decisión, tanto a nivel interno, como de los órganos de gobierno.

La aplicación de QCA provee de un marco analítico en el que se revelan las condiciones (o combinaciones de las mismas) necesarias y/o suficientes que permiten la consecución de los resultados.

La tabla 4.3 presenta el grado de pertenencia de los organismos objeto de estudios a los sets o conjuntos definidos para cada una de las tres condiciones (capacidad de aprendizaje, institucionalización de la función de evaluación y políticas y prácticas evaluativas) y de los dos resultados (calidad y uso sostenible).

En la tabla 4.3 se emplean abreviaturas, tanto para las condiciones como para los resultados: **CA** es capacidad de aprendizaje evaluativo, **INS** es el nivel de institucionalización de la función de evaluación, **PP** la existencia de políticas y prácticas evaluativas, **QUA** la calidad de la evaluación, y **USO** el uso sostenible de la evaluación. Cada condición se subdivide en tres indicadores reflejados en la tabla 4.3 utilizando la misma numeración identificativa utilizada en la tabla 4.2.

La determinación del grado de pertenencia a cada una de las condiciones y resultados se obtiene a partir de la asignación de valores a cada organismo (0, 0.3, 0.6, 1) que se corresponden con la puntuación resultante del proceso de medición de los indicadores establecidos para cada una de las condiciones arriba descritas, en base a los informes de *peer review* (revisión por pares) mencionados. Para definir la puntuación total por factor explicativo, se ha obtenido la media (promedio) de los resultados obtenidos en cada uno de los tres indicadores por variable y resultado.

**Tabla 4.3. Sumario de puntuación en escala difusa (fuzzy set)**

Agencias		CA	INS	PP	QUA	USO
UNO Habitat	1	0.33	0.33	0.66	0.66	0.66
	2	0.66	0.33	0.33	1.00	0.66
	3	1.00	0.66	0.66	0.66	1.00
		<b>0.66</b>	<b>0.33</b>	<b>0.66</b>	<b>0.66</b>	<b>0.66</b>
PNUD	1	1.00	0.33	0.33	0.33	1.00
	2	0.33	0.33	0.33	0.33	0.33
	3	0.66	0.33	0.33	0.66	1.00
		<b>0.66</b>	<b>0.33</b>	<b>0.33</b>	<b>0.33</b>	<b>0.66</b>
OIOS	1	0.33	1.00	0.66	0.66	0.66
	2	0.00	0.66	0.66	1.00	0.33
	3	0.33	0.66	0.66	0.66	0.33
		<b>0.33</b>	<b>0.66</b>	<b>0.66</b>	<b>0.66</b>	<b>0.33</b>
PNUMA	1	0.66	0.33	0.66	0.66	0.66
	2	0.66	0.33	0.66	0.66	0.66
	3	0.66	0.66	0.66	0.66	0.33
		<b>0.66</b>	<b>0.33</b>	<b>0.66</b>	<b>0.66</b>	<b>0.66</b>
UNICEF	1	0.33	0.33	0.00	0.33	0.33
	2	0.00	0.33	0.00	0.33	0.00
	3	0.66	0.33	0.33	0.33	0.33
		<b>0.33</b>	<b>0.33</b>	<b>0.00</b>	<b>0.33</b>	<b>0.33</b>
FAO	1	0.33	0.66	0.33	0.33	0.00
	2	0.33	0.66	0.66	0.33	0.33
	3	0.66	0.66	0.33	0.33	0.33
		<b>0.33</b>	<b>0.66</b>	<b>0.33</b>	<b>0.33</b>	<b>0.33</b>
ONUUDI	1	0.33	0.66	0.66	0.33	0.66
	2	0.66	0.33	0.66	0.33	0.33
	3	0.33	0.00	0.33	0.33	0.33
		<b>0.33</b>	<b>0.33</b>	<b>0.66</b>	<b>0.33</b>	<b>0.33</b>
FMAM	1	0.33	1.00	0.66	1.00	0.33
	2	0.00	1.00	1.00	0.33	0.33
	3	0.33	0.66	0.66	0.66	0.66
		<b>0.33</b>	<b>1.00</b>	<b>0.66</b>	<b>0.66</b>	<b>0.33</b>
PMA	1	0.33	0.66	0.33	0.33	0.33
	2	0.33	0.33	0.33	0.33	0.33
	3	0.00	0.33	0.33	0.33	0.00
		<b>0.33</b>	<b>0.33</b>	<b>0.33</b>	<b>0.33</b>	<b>0.33</b>
ONU Mujeres	1	0.66	0.66	1.00	0.66	0.33
	2	0.66	0.66	1.00	0.66	0.66
	3	1.00	0.33	0.66	1.00	1.00
		<b>0.66</b>	<b>0.66</b>	<b>0.66</b>	<b>0.66</b>	<b>0.66</b>

Todos los datos se han obtenido de los informes de revisión por pares llevados a cabo desde el 2006 hasta la actualidad. Puesto que algunos organismos (PMA, UNICEF, PNUD) realizaron su informe de revisión por pares hace ya algunos años es probable que la situación actual de dichos organismos frente a cada una de las condiciones y resultados haya evolucionado. En este sentido, es importante destacar que los datos del análisis del QCA no pretenden mostrar la situación más actual de las entidades objeto de estudio, sino el estado de la capacidad evaluativa en el momento cuando se realizó el informe de revisión.

Algunos organismos, (FMAM, PMA y PNUD) han realizado más de un ejercicio de revisión por pares, por ejemplo, FMAM realizó uno en el 2006, y otro en el 2014. En estos casos, se ha optado por mostrar en el fsQCA la situación de la función de evaluación en la entidad tal y como se presenta en el primer informe de revisión por pares y, a continuación, se ha analizado la evolución de la función de evaluación de dichas agencias entre los dos procesos de revisión por pares efectuados. Esta decisión se ha tomado principalmente por dos razones:

La primera es que los segundos informes de revisión por pares no presentan una información tan detallada de los criterios base para el análisis de la función de evaluación (credibilidad, utilidad e independencia) como los primeros informes, sino que inciden principalmente en aquellas áreas donde se encontraron ciertas debilidades en el primer informe realizado.

La segunda es que la consideración del primer informe permite captar la variabilidad y, por tanto, la evaluación de las funciones de evaluación a partir de dichos informes. En este sentido, el tomar como base para el análisis los datos del primer informe de revisión por pares permite analizar con más detalle la reacción y cambios efectuados en la capacidad de evaluación de dichas entidades tras el proceso de revisión, y tener una visión más detallada de la evolución de dichas agencias a lo largo del tiempo.

En base a lo anterior, se puede afirmar que la principal consideración para el análisis ha sido la consistencia temporal entre condiciones y resultados.

En este sentido, hay que señalar también que QCA es una técnica particularmente interesante en situaciones donde el tamaño de la muestra es medio y, por tanto, no se disponen de suficientes observaciones como para poder aplicar análisis estadísticos tradicionales (Ragin y Rihoux 2004; Rihoux y Grimm 2006). Para ello, QCA considera todas las combinaciones posibles de las condiciones de cara a obtener los resultados y se centra en aquellas condiciones (o combinaciones de ellas) que son plausibles para la persona investigadora experta, erigiéndose así como una técnica lógica frente a las técnicas estadísticas de naturaleza probabilística (González, López y Vidal 2012). En el caso que nos atañe, la muestra se compone de 10 organismos del Sistema de NU, todos ellos sometidos a un proceso de revisión por pares, lo que hace que QCA se convierta en una herramienta especialmente conveniente para comparar los diferentes casos.

Igualmente, frente a los métodos cualitativos, uno de los principales valores añadidos de QCA es que permite afrontar, de forma coherente, el hecho de encontrarnos ante muestras en el ámbito de las ciencias sociales caracterizadas por una diversidad limitada de casos, ya que no se reflejan empíricamente todas las combinaciones de factores posibles que provocan el resultado objeto de estudio. Es decir, cuando analizamos un fenómeno social con el objetivo de revelar una relación causal entre condiciones y resultados, QCA considera todas las combinaciones posibles de condiciones que pueden producirse y acude a la muestra para señalar, de entre los miembros de la misma, aquellos en los que se produce una u otra combinación de condiciones. La diversidad limitada frecuente en las ciencias sociales hace referencia a que encontramos múltiples combinaciones de condiciones o factores que no están representados empíricamente.

Por último, para preservar su capacidad de desenvolverse en un ámbito de relaciones causales, QCA ofrece una ruta para afrontar esta situación, de acuerdo al criterio del investigador/a, frente a esos casos en los que no tengamos respaldo empírico (también denominados contra-fácticos). Ragin (1987) propone tres posibilidades en función de la posición que adopta el/la investigador/a frente a aquellas combinaciones de condiciones no respaldadas empíricamente. Así, el/la investigador/a debe decidir si considera o no el resultado que en esas situaciones particulares se produce. Las tres rutas propuestas son: la solución compleja, la solución parsimoniosa (o más simple), y la intermedia.

La denominada **solución compleja** propone que frente a combinaciones de factores que no se ven reflejadas empíricamente (no hay casos), el/la investigador/a considere que el resultado no se produce. Responde, por tanto, a una posición más conservadora donde el investigador/a es reacio a ofrecer una relación causal entre condiciones y resultados basada, en buena medida, en casos no respaldados por la realidad. Se denomina solución compleja porque al haber menos rutas posibles para alcanzar los resultados, es más difícil simplificar entre las condiciones.

En contraste, la solución **parsimoniosa**, propone que el/la investigador/a considere que sí se produce el resultado cuando nos encontremos frente a los casos contrafácticos (no respaldados empíricamente). Lógicamente, responde a una posición más optimista y que ofrece más rutas posibles para alcanzar los resultados.

Por último, la solución **intermedia** propone que el/la investigador/a sea quién, caso por caso, determine si debe considerar que se produce o no el resultado en aquellas combinaciones de condiciones que no están respaldadas empíricamente. El criterio adoptado depende, en gran medida, de su conocimiento sobre cada caso particular (Ragin y Sonnet 2005).

De esta manera, QCA es un método que combina elementos cuantitativos con criterios cualitativos relativos a la experiencia y conocimientos del investigador/a, y que ofrece múltiples oportunidades para reflejar su experiencia y conocimiento a lo largo del proceso.

#### **4.4. Las entidades objeto de estudio**

En el anexo 1 se describen en detalle los diez organismos que forman parte de la muestra para el análisis de QCA enfatizando en las siguientes cuestiones:

- ✓ Mandato
- ✓ Marco programático
- ✓ Gobernabilidad
- ✓ Recursos
- ✓ Función de evaluación
- ✓ Proceso de revisión por pares

Como se ha mencionado anteriormente, estas entidades se han sometido a procesos de revisión por pares de sus funciones de evaluación, y responden a una muestra representativa de cómo se realiza la práctica evaluativa en NU. Los diez organismos incluyen Fondos y Programas, Organismos Especializados, Órganos de la Secretaría, y OIOS como entidad con la evaluación como mandato específico.

## **5. Análisis empírico: Resultados y discusión**

A continuación se presenta el análisis empírico a través del método QCA en sus diferentes secuencias lógicas para cada uno de los resultados del modelo. Se comenzará con el resultado de calidad, para proseguir posteriormente con el de uso sostenible.

En cada uno de los resultados, se realizará un análisis de las condiciones necesarias, a través de la comprobación de los índices de consistencia, y se procederá a su revisión gráfica en el diagrama XY. Posteriormente, se realizará un análisis de las condiciones suficientes a través de los resultados que se obtengan de la tabla de la verdad, y se describirán los caminos o combinaciones de condiciones que facilitan el logro de los resultados tanto para la calidad como para el uso.

### **5.1. La calidad de la evaluación**

La calidad es uno de los resultados esperados de la capacidad evaluativa en nuestro modelo. El método de QCA, una vez identificadas las condiciones que se consideran relevantes desde un punto de vista teórico para garantizar la calidad en las evaluaciones, procederá a analizar el grado de necesidad y suficiencia de las mismas.

Para determinar resultados positivos en cuanto a la calidad de las evaluaciones, se ha partido del análisis de los indicadores siguientes: 1) Calidad de los informes de evaluación, 2) existencia y aplicabilidad de controles de calidad, y 3) valoración positiva por parte de los usuarios intencionados de los informes de evaluación/información evaluativa generada en la organización. La información para la comprobación de dichos indicadores se ha obtenido de los informes de revisión por pares de las agencias objeto de estudio.

#### **5.1.1 Análisis de las condiciones necesarias**

Tal y como se ha mencionado anteriormente, QCA, en un marco de relaciones causales, compara los casos de una muestra de forma en que éstos se puedan agrupar entorno a dos categorías de condiciones: condiciones necesarias y condiciones suficientes.

La identificación de las condiciones necesarias, que es el primer paso, se realiza a través del 'método de concurrencia' mencionado anteriormente.

Se consideran así únicamente los elementos de la muestra (de entre los 10 organismos de las NU en nuestro caso) en los que se ha producido el resultado objeto de estudio (calidad de la evaluación), y se analizan las condiciones comunes en todos estos casos. La existencia de una condición necesaria requiere la presencia de la misma en todos aquellos casos donde se produzca el resultado. En términos de lógica difusa, esto implica que los valores de pertenencia de los casos de una muestra al conjunto

que representa la condición deben ser iguales o mayores que los valores de pertenencia al conjunto que representa el resultado.

El **índice de consistencia** permite identificar hasta qué punto las condiciones (bien de manera individual o combinada) pueden ser consideradas necesarias (Ragin 2006)<sup>77</sup>. Así, en una escala de 0 a 1, el índice de consistencia indica el grado de presencia de una condición cuando el resultado objeto de estudio se produce. En la literatura (Schneider y Wagemann 2007, Pérez 2007) se recomienda considerar que las condiciones sean necesarias únicamente en situaciones donde los índices de consistencia sean muy altos, si no máximos, especialmente ante muestras pequeñas, como ocurre en este caso. A continuación vamos a analizar en detalle los resultados del índice de consistencia y sus consecuencias en la interpretación en cuanto a la necesidad de las condiciones.

El **índice de cobertura** por su parte, indica también en una escala de 0 a 1, el grado de relevancia explicativa de una condición con respecto al resultado<sup>78</sup>. Es importante destacar que la cobertura solo se interpreta en el caso de resultados consistentes, es decir, el análisis del grado de cobertura en índices de consistencia bajos no se considera de utilidad (Schneider y Wagemann 2007).

La Tabla 5.1 presenta los resultados sobre el análisis de las condiciones necesarias para el resultado de 'calidad de la evaluación'.

**Tabla 5.1. Análisis de condiciones necesarias para el resultado 'calidad de la evaluación'**

Condición	Consistencia	Cobertura
CA	0,866	0,928
ca	0,806	0,741
INS	0,866	0,864
ins	0,737	0,724
PP	0,933	0,933
pp	0,676	0,663

Nota: La tabla 5.1 describe los índices de cobertura y consistencia de las condiciones objeto de estudio (CA, INS, y PP), y también de la ausencia de dichas condiciones (ca, ins y pp).

Si determinamos un índice de consistencia máximo de 1, estaríamos estableciendo la premisa de que para que una condición sea necesaria, no puede existir un solo caso donde dicha condición no esté presente y se produzca el resultado. Los resultados de la Tabla 5.1 desde un prisma de consistencia máximo (índice 1) permiten afirmar que no existen condiciones necesarias a nivel individual para conseguir altos niveles de calidad en la evaluación. Es decir, si analizamos una a una las condiciones a nivel individual, nos encontramos que, para cada una de ellas, al menos se produce una

<sup>77</sup> Así, en términos de Ragin (2006, pág. 291), la consistencia analiza el grado de aproximación de un sub-conjunto causal al resultado.

<sup>78</sup> Así, la cobertura se centra en la relevancia explicativa de un sub-conjunto con respecto al resultado (Ragin 2006, pág. 291).

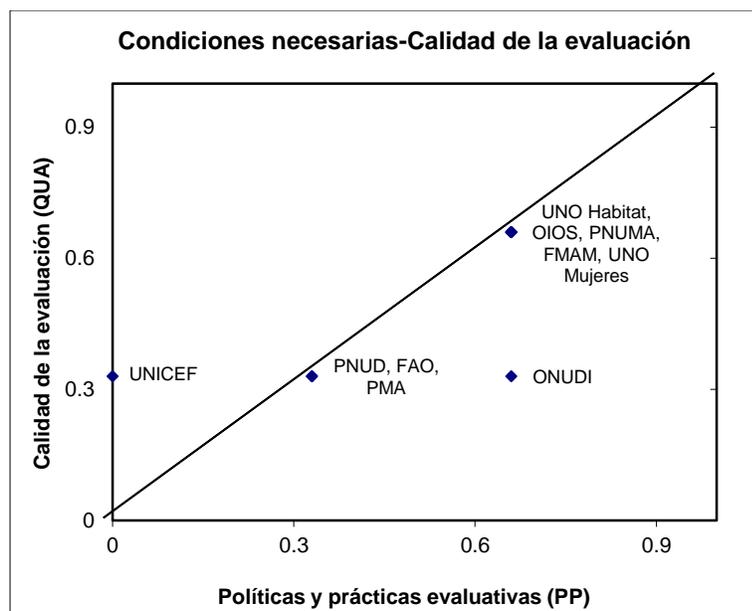
situación donde el valor de pertenencia de la condición al subconjunto es menor al valor de pertenencia al resultado. Esto implica que, en las tres condiciones objeto de estudio (CA, INS y PP), existe al menos un caso, donde los valores resultantes de estas condiciones no están al mismo nivel que los niveles de calidad, sino que son inferiores. De esto subyace, *a priori*, que la presencia de las condiciones, no es necesaria para el logro de la calidad.

Sin embargo, si establecemos un índice de consistencia de 0,90, nos encontramos ante la situación de que las políticas y prácticas evaluativas pueden constituirse en una condición necesaria (puesto que su índice de consistencia es de 0,933). Ante esta disyuntiva, el análisis gráfico es de fundamental importancia para poder determinar o descartar la existencia de casos irrelevantes que puedan estar influyendo en los valores de consistencia (Schneider y Wagemann 2007). En relación al índice de cobertura, podemos observar que las tres condiciones seleccionadas son altamente relevantes para explicar el resultado de calidad.

Por lo tanto, el análisis de la necesidad de las condiciones se ha complementado gráficamente para todas las condiciones. Para que se considere una condición como necesaria, *“en principio todos los casos deben aparecer sobre o debajo de la línea divisoria”* (Ragin 2000, pág. 215).

Los gráficos 5.1, 5.2 y 5.3 muestran la relación entre cada una de las condiciones y el resultado calidad de la evaluación.

**Gráfico 5.1. Políticas/Prácticas Evaluativas y Calidad de la Evaluación**

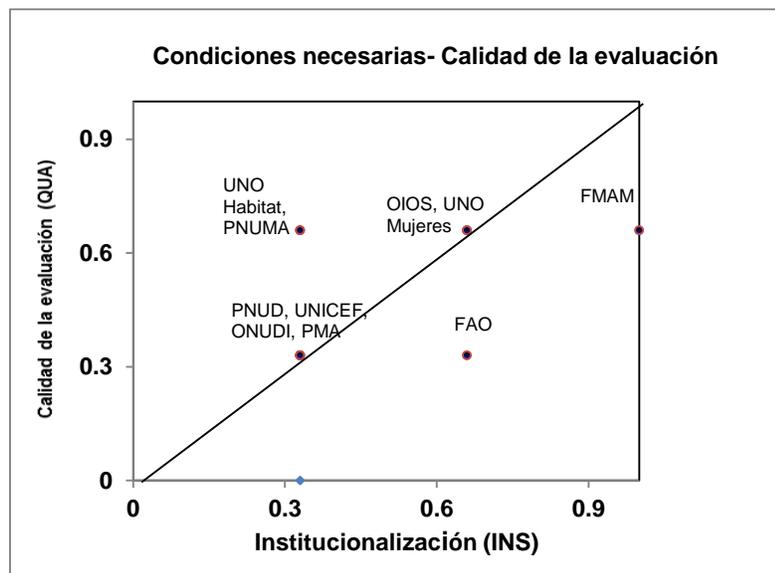


En el Gráfico 5.1, los diez organismos se agrupan en torno a cuatro puntos del diagrama. Un total de ocho de las diez agencias se sitúan sobre los puntos de la línea divisoria, lo que indica que, en estas ocho agencias, el grado de pertenencia de la condición al conjunto es igual al grado de pertenencia al resultado. Así, las agencias PNUD, FAO y PMA tienen en común bajos niveles de políticas y prácticas evaluativas (0.33) y bajos resultados respecto a la calidad de sus evaluaciones (0.33). Por su parte, se observan altos niveles de consolidación de políticas y prácticas evaluativas en las agencias UNO Habitat, OIOS, PNUMA, FMAM y UNO Mujeres y, a su vez, altos niveles de calidad encontrados en sus evaluaciones (0.66 en ambos casos).

En el caso de UNICEF, el Gráfico 5.1 muestra que no existe grado de pertenencia a la condición de Prácticas y Políticas evaluativas (0,00), y los resultados en calidad, a pesar de ser más altos (0,33), no llegan al nivel mínimo de pertenencia al resultado (0,5). En este sentido, cabe destacar que, cuando UNICEF se sometió al proceso de revisión por pares (2006), todavía no contaba con una Política de Evaluación, ni procesos evaluativos sólidos (tanto en la elaboración del plan de evaluación como en el proceso de seguimiento a las recomendaciones). El caso de UNICEF es el único caso que impide que la condición de políticas y prácticas evaluativas pueda ser considerada necesaria. Sin embargo, a través del gráfico 5.1 vemos que, en realidad, se trata de un caso considerado como irrelevante, puesto que los resultados en calidad no llegan al nivel mínimo de pertenencia al resultado (0,5). Por lo tanto, si descartamos el caso de UNICEF, nos encontramos con que la condición PP (políticas y prácticas evaluativas) sí es condición necesaria para el logro de la calidad en la evaluación. Esto es debido a que no existe un solo caso en nuestra muestra donde contando con valores que superen el umbral de pertenencia (0,5) a políticas y prácticas evaluativas, no se consigan buenos resultados en calidad.

Finalmente, ONUDI por su parte, despliega una pertenencia elevada al conjunto de Prácticas y Políticas evaluativas que no se corresponde con la calidad de las evaluaciones. Así, este caso pone de manifiesto que la existencia de políticas y prácticas evaluativas (0,66) no garantiza resultados en calidad (0,33) y, por tanto, sugiere cuestionarse la presencia de políticas y prácticas evaluativas (PP) como una condición suficiente por sí misma para el logro de resultados positivos a nivel de calidad (QUA). Este resultado será confirmado posteriormente en el análisis de suficiencia.

**Gráfico 5.2. Institucionalización y Calidad de la Evaluación**



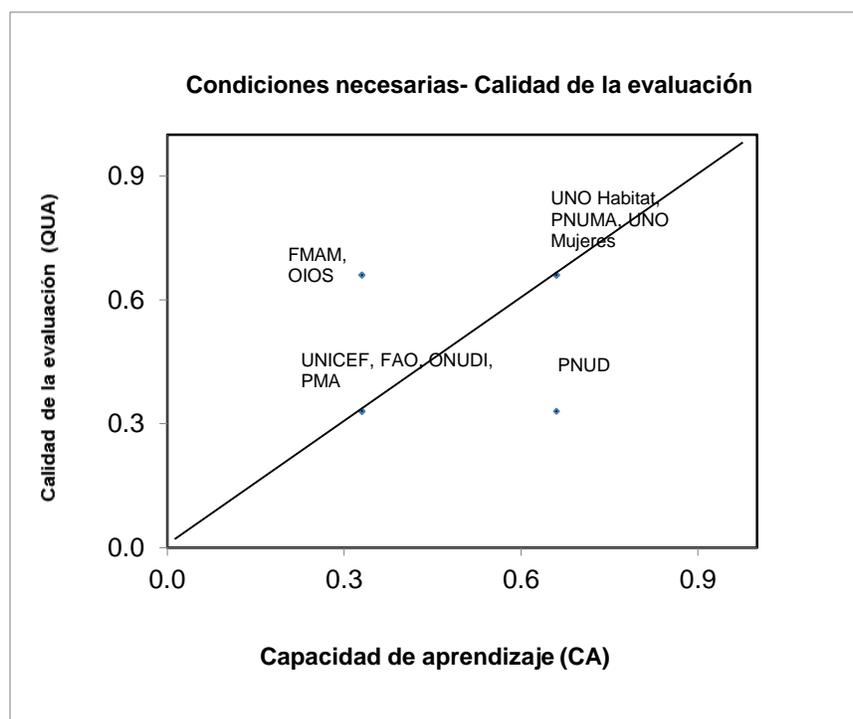
En el Gráfico 5.2, los diez organismos se agrupan en torno a cinco puntos del diagrama. Seis de las diez agencias se sitúan sobre la línea divisoria, donde su grado de pertenencia de la condición al conjunto es igual que el grado de pertenencia al resultado. En concreto, OIOS y UNO Mujeres coinciden en altos niveles de institucionalización de la función de evaluación (0.66) y, por tanto, independencia estructural y presupuestaria y altos niveles de calidad de sus evaluaciones (0.66). En el otro extremo, PNUD, UNICEF, ONUDI y PMA rebelan débiles niveles de institucionalización de su función de evaluación (0,33), correspondidos con débiles niveles en cuanto a calidad (0,33)<sup>79</sup>.

A su vez, el Gráfico 5.2 muestra dos casos (UNO Habitat, PNUMA) donde, a pesar de que los niveles de institucionalización son bajos (0,33), los niveles de calidad son importantes (0,66). Por tanto, estos dos casos descartan la condición de institucionalización como condición necesaria para el logro de resultados en calidad.

También se encuentra el caso contrario en el gráfico 5.2. Así, la función de evaluación de FAO está muy consolidada institucionalmente (0,66) pero, aun así, se observan ciertas debilidades en la calidad de sus informes (0,33). Por lo tanto, este resultado induce a pensar que la institucionalización no tiene por qué resultar una condición suficiente para el logro de la calidad.

<sup>79</sup> Las principales debilidades encontradas durante el proceso de revisión por pares con respecto a la institucionalización de la función de evaluación en dichos organismos responden a limitaciones de comunicación y reporte de la persona responsable de la unidad de evaluación a los órganos de gobierno, dificultades en la predictibilidad y transparencia presupuestaria en la asignación de recursos para evaluación y falta de recursos humanos para gestionar de manera consistente las evaluaciones y prestar el apoyo necesario a las oficinas de campo.

**Gráfico 5.3. Capacidad de Aprendizaje y Calidad de la Evaluación**



El Gráfico 5.3 muestra la diversidad de los organismos con respecto a su capacidad de aprendizaje, y la relación existente con sus resultados con respecto a la calidad de las evaluaciones. Un total de siete agencias se encuentran sobre la línea divisoria, donde su grado de pertenencia de la condición al conjunto coincide con el grado de pertenencia al resultado. UNO Habitat, PNUMA y UNO Mujeres se encuentran entre las agencias que cuentan con una sólida capacidad de aprendizaje y sólidos resultados en cuanto a calidad de sus evaluaciones (0,66 en ambos casos). UNICEF, FAO, ONUDI y PMA presentaron niveles algo más débiles en su capacidad de aprendizaje en el momento de la revisión por pares, que se corresponden con su nivel de calidad (0,33 en ambos casos)<sup>80</sup>.

Al igual que en el análisis de las anteriores condiciones, el Gráfico 5.3 también contempla casos que cuestionan que la capacidad de aprendizaje sea, en sí misma, necesaria para lograr resultados en calidad (OIOS, FMAM). Así, en estos dos casos, los resultados de calidad son positivos (0,66), a pesar de contar con niveles bajos en cuanto a su capacidad de aprendizaje (0,33). Los resultados de estas dos agencias, por tanto, descartan que la capacidad de aprendizaje sea condición necesaria para lograr resultados en calidad.

<sup>80</sup> Entre las debilidades encontradas en las agencias en cuanto a su capacidad de aprendizaje cabe destacar la falta de una estrategia clara de diseminación de la información evaluativa, cierta inconsistencia en los niveles de participación y consulta en las evaluaciones, y confusión a nivel de organización sobre lo que la evaluación es, y sobre lo que puede aportar con respecto a otras formas de análisis (i.e, monitoreo).

Por su parte, el PNUD muestra resultados positivos en cuanto a su capacidad de aprendizaje (0,66), que no conllevan necesariamente a los mismos resultados en cuanto a su calidad (0,33), lo que cuestiona la capacidad de aprendizaje como condición suficiente para lograr resultados positivos en la calidad de las evaluaciones.

Por lo tanto, los resultados combinados del índice de consistencia y del análisis gráfico concluyen que la existencia de políticas y prácticas evaluativas (PP) es condición necesaria para el logro de calidad en la evaluación. Sin embargo, los resultados también dejan entrever que no es condición suficiente.

Lograr que en las organizaciones se realicen evaluaciones de calidad depende de muchos factores. Un factor clave es la existencia de políticas y prácticas evaluativas, pero no el único. El análisis de suficiencia más adelante analizará los diferentes caminos (o combinación de condiciones) que llevan a las organizaciones a resultados de calidad. Teniendo en cuenta que, en nuestro análisis, la existencia de políticas y prácticas evaluativas (PP) es condición necesaria para lograr resultados en calidad, esta condición deberá estar presente en todos y cada uno de los caminos posibles de suficiencia.

Los resultados del análisis de las condiciones necesarias permiten concluir que la calidad es un aspecto multidisciplinar al depender de varias variables que van más allá de la puesta en marcha de políticas y prácticas sólidas que faciliten el rigor metodológico y la aplicación de procedimientos evaluativos. A pesar de que la presencia de políticas y prácticas evaluativas es fundamental para lograr resultados positivos en calidad, dicha presencia debe ir acompañada de otras medidas para garantizar la calidad en las evaluaciones como, por ejemplo, la dotación adecuada de recursos financieros y personal, así como el garantizar que la función de evaluación sea independiente de presiones internas o externas.

Combinar prácticas y políticas evaluativas sólidas con otras medidas es especialmente relevante en entidades donde la función de evaluación está descentralizada. De las entidades objeto de estudio, en cuatro de ellas el trabajo evaluativo se produce en su mayoría a nivel descentralizado (UNICEF, FAO, PMA, y UNO Mujeres). Las conclusiones de los estudios de revisión por pares, especialmente para el caso de UNICEF, FAO y PMA, coinciden en observar importantes debilidades en las funciones de evaluación descentralizadas con negativos resultados en cuanto a la calidad de esas evaluaciones. Entre sus principales carencias, además de falta de recursos financieros y técnicos para llevar a cabo actividades evaluativas a nivel descentralizado, también se observa que las políticas de evaluación no clarifican suficientemente los roles y nivel de apoyo que la unidad de evaluación central debe proporcionar a las oficinas regionales y nacionales<sup>81</sup>.

---

<sup>81</sup> En el caso de PMA, la segunda revisión por pares efectuada en 2014 siguió observando serias debilidades en la calidad de la evaluación descentralizada y propuso tres modelos alternativos (un modelo donde la evaluación es totalmente centralizada, y dos mixtos), para tratar de mejorar la situación existente (UNEG/WFP 2004, pág. 19-20).

La situación es algo diferente en el caso de ONU Mujeres. El panel de revisión por pares no detecta diferencias notables en cuanto a la calidad de las evaluaciones realizadas a nivel descentralizado y valora positivamente la puesta en marcha de controles de calidad rigurosos para el monitoreo de las evaluaciones descentralizadas así como la claridad de roles y responsabilidades entre las oficinas de campo y la Oficina de Evaluación (UNEG/UNWomen 2014, págs. 7, 14).

### 5.1.2 Análisis de las condiciones suficientes

La identificación de condiciones causales suficientes se realiza siguiendo las siguientes etapas:

- ✓ Identificación de todas las combinaciones de condiciones causales posibles.
- ✓ Análisis de la distribución de los casos en función de las combinaciones identificadas.
- ✓ Detección de las condiciones suficientes a partir del criterio de consistencia.

Se considera que, para que una condición sea suficiente, el resultado debe producirse siempre que dicha condición esté presente (aunque también pueda producirse cuando esta condición no esté presente). El análisis de las condiciones suficientes es realizado por fsQCA a través del algoritmo de la denominada 'tabla de la verdad'<sup>82</sup> (*truth table*) de Quine McClusky (Schneider y Wagemann 2007). A partir de este análisis, se observa que una combinación de condiciones es suficiente para que se produzca el resultado cuando el grado de pertenencia de cada caso (organismo) al conjunto representado por la combinación de condiciones es igual o menor que su pertenencia al resultado. El índice de consistencia de cada combinación es el que nos indicará así el grado de suficiencia. A continuación se presenta la tabla de la verdad para el resultado 'calidad de la evaluación'<sup>83</sup>.

---

<sup>82</sup> La tabla de la verdad representa cada combinación lógica de los valores de las condiciones causales. Por tanto, en la tabla de la verdad hay tantas filas como combinaciones lógicas posibles. La tabla de la verdad, una vez realizada permite la simplificación de las configuraciones causales, y facilita la búsqueda de patrones comunes.

<sup>83</sup> El número de combinaciones causales es una función exponencial de las condiciones causales del modelo  $2^K$ , donde k es el número de combinaciones causales (Ragin 2008, pág. 24). En nuestro caso, puesto que tenemos tres condiciones, tenemos un total de ocho combinaciones posibles. Sin embargo, existe un caso posible que no está recogido en nuestra muestra, el contra fáctico en la combinación 1, 1. 0. Es, por ello, que tenemos un total de 7 filas.

**Tabla 5.2. Tabla de la verdad sobre el resultado ‘Calidad de la Evaluación’**

CA	INS	PP	Agencias	Casos	QUA	Consistencia
1	1	1	UNO Mujeres	1	1	1,000
0	1	1	OIOS, FMAM	2	1	1,000
1	0	1	UNO Habitat, PNUMA	2	1	1,000
0	1	0	FAO	1	0	0,909
1	0	0	PNUD	1	0	0,900
0	0	1	UNIDO	1	0	0,890
0	0	0	UNICEF, PMA	2	0	0,809

Nota: El umbral de consistencia para las condiciones suficientes se ha establecido en 0.95<sup>84</sup>.

En nuestro análisis existe una combinación de condiciones que no está representada empíricamente, es decir, no tenemos un organismo en el cual se presente esta determinada combinación dentro de las ocho combinaciones posibles (*remainder o contra fáctico*). En concreto esta combinación es 1 1 0, que representa: alta capacidad de aprendizaje (CA), alta institucionalización (INS) y bajos niveles de políticas y prácticas evaluativas (pp).

Este caso, sin respaldo empírico, se ha tratado tal y como sugieren Ragin y Sonnet (2005), analizando el balance de aplicar un enfoque de parsimonia (donde el/la investigador/a considera este resultado como positivo), un enfoque de complejidad (donde el resultado se considera negativo), o un enfoque intermedio (se decide para cada contra fáctico si debe considerarse el resultado positivo o negativo en función de su conocimiento sobre los casos). Partiendo de la base de que no existe en el marco teórico una lógica definida que prescriba un resultado plausible ante esta combinación, se ha optado por la opción de presentar todos los diferentes enfoques posibles.

Así, el análisis de la tabla de la verdad ofrece dos resultados posibles en función del tratamiento a los contra fácticos. En nuestro caso, solo existe un contra fáctico (la combinación 1, 1, 0, explicada anteriormente) y, por lo tanto, puesto que solo existe un contra fáctico, la solución compleja es igual a la intermedia (ausencia de resultado). Por el contrario, la solución parsimoniosa representa la presencia del resultado ante dicha combinación. A continuación, se muestran los resultados del análisis de las condiciones suficientes del resultado ‘calidad de la evaluación’ derivado de la tabla de la verdad (Tabla 5.3).

<sup>84</sup>Schneider y Wagemann (2012) recomiendan establecer un umbral un poco más alto para las condiciones necesarias que para las condiciones suficientes, debido a la mayor rotundidad de su significado. Puesto que se ha establecido un índice máximo de consistencia para las condiciones necesarias, para las condiciones de suficiencia se ha considerado un 0,95.

**Tabla 5.3. Análisis de condiciones suficientes en ‘Calidad de la Evaluación’.**

Solución	Compleja		Parsimoniosa	
	CA*PP	PP*INS	CA*PP	PP*INS
<b>Agencias</b>	UNO Habitat PNUMA UNO Mujeres	OIOS FMAM UNO Mujeres	UNO Habitat PNUMA UNO Mujeres	OIOS FMAM UNO Mujeres
<b>Consistencia</b>	1,000	0,921	1,000	0,921
<b>Cobertura total</b>	0,800	0,802	0,800	0,802
<b>Cobertura exclusiva</b>	0,133	0,266	0,133	0,266
<b>Consistencia de la solución</b>	0,935		0,935	
<b>Cobertura de la solución</b>	1,000		1,000	

En nuestro caso, las dos soluciones (la compleja y la parsimoniosa) obtienen los mismos niveles de consistencia y cobertura puesto que ofrecen los mismos caminos, por lo que el contra fáctico no altera los resultados (0,935 y 1,000 respectivamente). Por su parte, los índices de cobertura exclusiva indican el grado de explicación única o ‘exclusiva’ de dichas configuraciones de condiciones, que en el caso de PP\*INS es algo superior que en el caso CA\*PP, lo que induce a pensar que ante resultados satisfactorios de calidad, existe un mayor número de casos explicados por la combinación PP\*INS. La tabla de la verdad identifica dos ‘caminos’, o condiciones suficientes que llevan a altos niveles de calidad de la evaluación. Estos son:

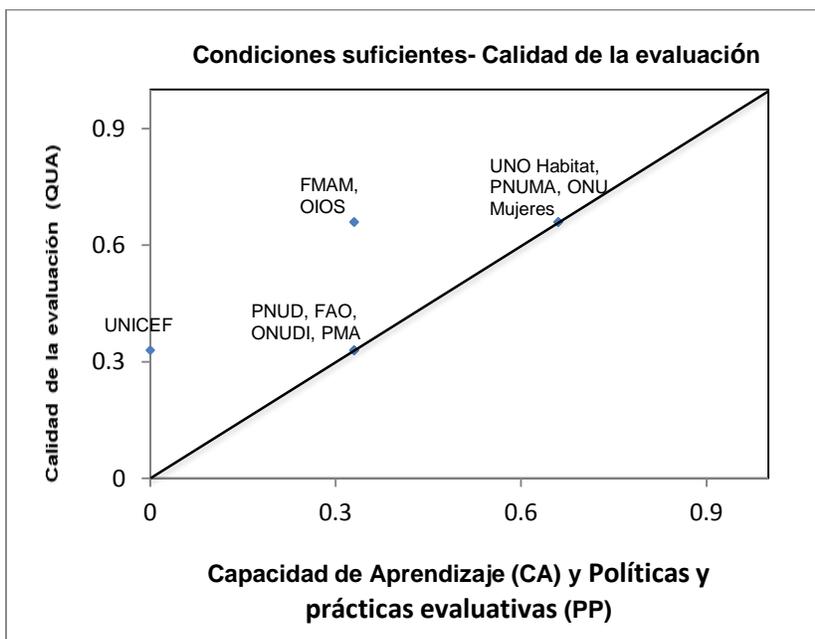
Condición suficiente 1: Alta capacidad de aprendizaje (CA) combinada con alto nivel de políticas y prácticas evaluativas (PP)

Condición suficiente 2: Alto nivel de institucionalización de la evaluación (INS) combinado con alto nivel de políticas y prácticas evaluativas (PP)

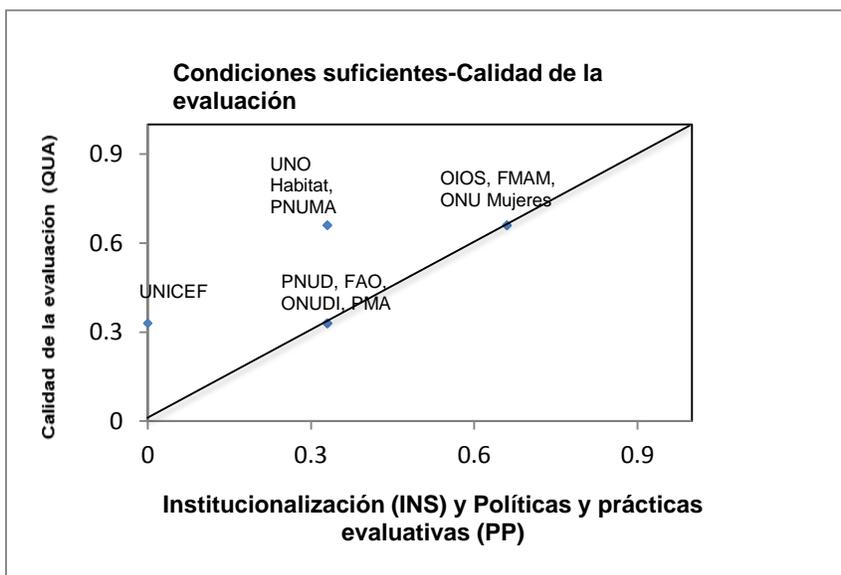
Los caminos que ofrece la tabla de la verdad confirman la condición de políticas y prácticas evaluativas como condición necesaria, puesto que esta condición se encuentra presente en las dos soluciones.

Las figuras 5.4 y 5.5 presentan los mismos resultados gráficamente. Para que una condición sea suficiente, “*todos los casos deben estar situados alrededor o sobre la línea divisoria*” (Ragin 2000, pág. 236).

**Gráfico 5.4. La calidad de la evaluación desde el aprendizaje y las políticas/prácticas evaluativas**



**Gráfico 5.5. La calidad de la evaluación desde la institucionalización y las políticas/prácticas evaluativas**



Así, la primera combinación de condiciones suficientes responde a altos niveles en la capacidad de aprendizaje (CA) junto con altos niveles de políticas y prácticas evaluativas (PP). La segunda combinación de condiciones suficientes responde a altos

niveles de institucionalización (INS) junto con altos niveles de políticas y prácticas evaluativas (PP). Las dos combinaciones son muy consistentes con valores de 0,935. Por su parte, los niveles de cobertura son máximos (1,000), lo que confirma su total relevancia explicativa en los casos donde existe alta calidad.

De estos resultados del análisis de suficiencia en la calidad de las evaluaciones se desprenden las siguientes conclusiones:

- ✓ La condición de políticas y prácticas evaluativas (PP) se confirma como condición necesaria para poder obtener positivos resultados en cuanto a la calidad de las evaluaciones.
- ✓ Se confirma que no existen condiciones que, a nivel individual, sean suficientes por si mismas para lograr objetivos de calidad.
- ✓ Las políticas y prácticas evaluativas (PP) son, por si solas, condición no suficiente para el logro de la calidad, si bien actúan como pivote para el logro de buenos resultados cuando van acompañadas de: 1) una sólida capacidad de aprendizaje (CA), o 2) unos niveles consolidados de institucionalización de la función de evaluación (INS).
- ✓ Los resultados engloban a las agencias objeto de estudio en dos grupos, que principalmente están divididos según la estrategia adoptada por las mismas en cuanto a la priorización otorgada a la capacidad de aprendizaje o al nivel de institucionalización de la función de evaluación.
- ✓ Así, en un primer grupo se encuentran entidades como UNO Habitat y PNUMA, agencias que han enfocado sus esfuerzos en fortalecer la cultura de aprendizaje. a la vez que establecen políticas y prácticas evaluativas alineadas con los estándares UNEG. En un segundo grupo, están OIOS y FMAM, que han optado por priorizar ante todo el fortalecimiento institucional de la función de evaluación, aunque esto haya supuesto cierto aislamiento y distancia con los enfoques participativos y, en general, con los elementos inherentes a la cultura de aprendizaje. Por último, UNO Mujeres está presente en los 2 grupos, puesto que es el único caso de la muestra que ha obtenido valores consolidados en las tres condiciones.

Por la importancia que estos resultados tienen para diferenciar diferentes estrategias orientadas a fortalecer la calidad de las evaluaciones, en el siguiente sub-apartado se explican más en detalle las características de cada uno de los dos grupos arriba mencionados:

## **Grupo 1: Calidad evaluativa desde el aprendizaje y las políticas evaluativas**

UNO Habitat, PNUMA y UNO Mujeres son tres agencias que coinciden en el logro de altos niveles de calidad en las evaluaciones, y en la importancia otorgada desde sus órganos de gerencia a la participación y, en general, a la promoción de la cultura del aprendizaje a través de la evaluación así como por el desarrollo de políticas y prácticas evaluativas sólidas.

Si desgranamos elementos comunes en dichas agencias (UNO Habitat, PNUMA y UNO Mujeres) en cuanto a su capacidad de aprendizaje (CA), las tres organizaciones cuentan con altos niveles de satisfacción y conocimiento acerca de los potenciales usos de la evaluación en la organización. El personal percibe así que la alta gerencia considera relevante la práctica evaluativa y esto se refleja a través de sus decisiones. Por citar algunos ejemplos de decisiones gerenciales particularmente apreciadas por el personal de dichas agencias, cabe destacar la introducción del Plan Estratégico e Institucional de Medio Término (MSTIP) en UNO Habitat, adoptado por la resolución 21/2 de Abril de 2007 donde existe un área denominada innovación que permite capturar conocimiento y lecciones aprendidas de la evaluación en el ciclo programático. Las entrevistas en el informe de revisión por pares reflejan que esta herramienta ha contribuido a realzar el propósito del aprendizaje con respecto al de la rendición de cuentas en relación con el trabajo en evaluación (UNEG/UN Habitat 2012, págs. 13 y 29).

Por su parte, la gerencia de PNUMA ha incorporado información estadística que refleja el estado de las recomendaciones de los informes de evaluación como parte del análisis de valoración del rendimiento de las direcciones técnicas (UNEG/UNEP 2012, pág. 16). En el caso de UNO Mujeres, es visible y apreciado por sus donantes así como por el personal, tanto en las oficinas de campo, como en las oficinas regionales y en la sede central, el apoyo de la gerencia al liderazgo de la función de evaluación de UNO Mujeres en UNEG y en otras redes de promoción de la práctica evaluativa como EvalPartners (UNEG/UN Women 2014, pág. 3).

Las tres agencias a su vez comparten amplios niveles de consulta y participación de actores clave en el proceso de evaluación, aunque en algunos casos, especialmente en ONU Habitat y PNUMA, todavía se podría incrementar la implicación de los países y las oficinas de campo especialmente en la planificación de las evaluaciones y la diseminación de hallazgos (UNEG/UN Habitat 2012, págs., 19, 25, UNEG/UNEP 2012, pág. 25).

En cuanto a actividades referentes a la diseminación de la información evaluativa, cabe destacar los esfuerzos puestos en marcha por UNO Mujeres. Dado que cerca del 90% de la evaluación se realiza de manera descentralizada, esta agencia apuesta por la diseminación de las evaluaciones a través de su portal web y de las plataformas de conocimiento, como los sistemas GATE y GERAAS<sup>85</sup>. Las síntesis de evaluaciones son

---

<sup>85</sup>GATE: Global Accountability and Tracking of Evaluation System. GERAAS: Global Evaluation Reports Assessment and Analysis System.

publicadas anualmente y forman parte del Informe Anual de Evaluación presentado por la dirección de la Oficina de Evaluación a la Junta Ejecutiva. A su vez desde 2014 se publican estudios de meta-evaluaciones. El acceso estadístico al estado de las recomendaciones es público y es parte de la estrategia de comunicación (UNEG/UN Women 2014, págs. 4, 22-23).

En referencia a las políticas y prácticas evaluativas, las tres agencias presentan una sólida aplicación de las mismas tanto en la política de evaluación, como en el plan de evaluación y el sistema de seguimiento a las recomendaciones de los informes. Así, la política de evaluación de UNO Mujeres fue aprobada por su Junta Ejecutiva en 2012 en línea con las Normas y Estándares UNEG. En el caso de PNUMA, la política de evaluación fue formalmente aprobada en 2009 por su Dirección Ejecutiva aunque no fue sometida a aprobación del Consejo Ejecutivo<sup>86</sup>. De los tres organismos, la política de evaluación de UNO Habitat es la que ha presentado un desarrollo menos acelerado. Hasta el 2013, en vez de política de evaluación en UNO Habitat existía un manual y unas guías sobre monitoreo y evaluación que cumplían las funciones de política, aunque carecían de algunos elementos básicos, como una clara y precisa definición de roles y responsabilidades para las tareas de planificación, implementación y gestión de las evaluaciones (UNEG Habitat 2012, págs.11 y 16). La política de evaluación finalmente se formula en el 2013 y efectivamente contempla un sólido alineamiento con los estándares UNEG.

El plan de evaluación en los tres casos se establece en base a varios criterios entre los que se incluye la relevancia estratégica y el valor innovador de las evaluaciones. Su elaboración contempla un proceso consultivo que incluye a órganos de gobierno, personal y donantes. En el caso de UNO Habitat y PNUMA, la aprobación se produce a nivel interno (UNEG Habitat 2012, pág. 11; UNEG UNEP 2012, págs. 19 y 30) y en el caso de ONU Mujeres el plan se presenta a la Junta Ejecutiva para su aprobación (UNEG/UN Women 2014, págs. 7, 20).

En las tres agencias existe un sistema de seguimiento de las recomendaciones de los informes de evaluación. Así, las diferentes unidades técnicas reportan periódicamente acerca del estado de las recomendaciones aceptadas a la unidad de evaluación. En el caso de PNUMA y UNO Habitat, la unidad de evaluación prepara un análisis de los resultados obtenidos que es sometido a comités internos de gerencia con el fin de que éstos puedan ser usados en la formulación del nuevo ciclo (UNEG/Habitat 2012, págs. 14 y 19; UNEG/UNEP 2012, págs. 23 y 33). UNO Mujeres da un paso más en cuanto a transparencia se refiere y elabora una base de datos sobre el estado de las recomendaciones que es pública (GATE) y por tanto accesible no solo por el personal interno sino también por sus órganos de gobierno, donantes y otros usuarios potenciales<sup>87</sup>.

---

<sup>86</sup>Según el panel de revisión por pares, el hecho de que el Consejo Ejecutivo no haya aprobado formalmente la Política de Evaluación, puede haber influenciado negativamente a la hora de establecer unos niveles de comunicación fluidos entre la Oficina de Evaluación y los órganos de gobierno (UNEG UNEP 2012, pág. 20).

<sup>87</sup><http://gate.unwomen.org/>

Si nos centramos en los niveles de calidad de la evaluación, se puede observar que las tres agencias coinciden en el establecimiento y aplicación de controles de calidad rigurosos en todo el proceso evaluativo, incluyendo la realización de informes de evaluación disponibles on-line. También los tres organismos despliegan un alto grado de satisfacción de su personal y alta gerencia en relación con la calidad de los informes de evaluación y en general con el trabajo desempeñado desde las unidades de evaluación (UNEG/UN Habitat 2012, págs. 14, 20 y 23; UNEG/UNEP 2012, págs. 23-27, UNEG/UN Women 2014, págs. 19-20).

Cabe destacar que dentro del grupo de estas tres agencias, también existen algunas diferencias especialmente en el grado de institucionalización de la función de evaluación y, sin embargo, estas diferencias no afectan el grado de calidad de la evaluación. Así, si nos centramos en el grado de independencia estructural podemos observar que, tanto en el caso de PNUMA como en el de UNO Habitat, ambas presentan unos niveles de independencia estructural algo más débiles que UNO Mujeres, ya que sus unidades de evaluación (PNUMA y UNO Habitat) no tienen canal de comunicación directo con sus órganos de gobierno (solo a través de la oficina del Director Ejecutivo<sup>88</sup>) y, por tanto, las oportunidades de interactuar con los órganos de gobierno son más limitadas y principalmente de carácter *ad hoc* (UNEG/Habitat 2012, Págs. 16 y 20; UNEG/UNEP2012 pág. 20).

Algo similar ocurre en el ámbito de los recursos financieros. El nivel de independencia y predictibilidad financiera de PNUMA y UNO Habitat es más débil<sup>89</sup> que en el caso de UNO Mujeres. A diferencia de PNUMA y UNO Habitat, en UNO Mujeres el presupuesto propuesto por la unidad de evaluación es aprobado directamente por los órganos de gobierno sin modificaciones y es considerado apropiado y suficiente para hacer frente al plan de evaluación (UNEG/UN Women 2014, págs. 12, 18).

Las diferencias en cuanto al grado de institucionalización no han afectado los resultados positivos en la calidad de las evaluaciones de dichos organismos, debido en gran parte a los esfuerzos destinados de forma combinada al fortalecimiento de la cultura de aprendizaje y a la aplicación de procedimientos evaluativos. Se deduce así que la cercanía y percepción positiva del personal con respecto a la función de evaluación influyen positivamente en la aplicación (no solo el establecimiento) de procedimientos evaluativos que inciden efectivamente en altos niveles de calidad. Esto indica que el grado de institucionalización puede ser un factor no tan importante para

---

<sup>88</sup>En el caso de UNO Habitat la política de evaluación del 2013 si contempla un canal de comunicación directo, pero no anteriormente (fecha de la realización de la revisión por pares).

<sup>89</sup>En UNO Habitat, la asignación de recursos para evaluación tiene ciertas limitaciones en cuanto a transparencia e independencia. Los datos no son fácilmente accesibles y dado que monitoreo y evaluación están juntos presupuestariamente, es difícil saber que parte está asignada a evaluación. Se considera que la parte de la evaluación esta marginada con respecto a la parte de monitoreo, a la que se dedica la mayoría del tiempo y los recursos. La unidad no tiene control sobre el presupuesto para evaluación. La implementación del plan de evaluación no está asegurada financieramente (UNEG/Habitat 2012, págs. v, vi, 19). En PNUMA, el presupuesto regular solo cubre salarios de personal y gastos administrativos y viajes. Solo representa el 50% del presupuesto destinado a evaluar proyectos, que está bajo el control de la gerencia de los programas, y las cifras no son fácilmente accesibles. Esto limita completamente la maniobrabilidad de la unidad de evaluación (UNEG/UNEP 2012 págs. v, vi, 17 y 18).

lograr resultados de calidad, si en su lugar existen de manera efectiva las otras dos condiciones analizadas: capacidad de aprendizaje (CA) y políticas y prácticas evaluativas (PP).

Así, en una posición contrapuesta a la arriba descrita nos encontramos a la FAO, la cual posee una función de evaluación institucionalmente sólida tanto a nivel de estructura como de recursos, pero que, a fecha del informe de la revisión por pares (Junio 2012), aún presentaba ciertas debilidades en cuanto a los niveles de cultura de aprendizaje y aplicación de políticas evaluativas. El Plan Inmediato de Acción aprobado por la Conferencia de la FAO en Junio de 2008 supuso la puesta en marcha de importantes esfuerzos para el fortalecimiento de la función de evaluación. Sin embargo, la aplicación de políticas evaluativas es un proceso gradual y que lleva tiempo. Así, los todavía incipientes resultados alcanzados en cuanto a calidad y uso de la evaluación, llevaron al panel que efectuó la revisión por pares de la FAO en 2012 a realizar recomendaciones dirigidas a fortalecer la correcta aplicación de procedimientos evaluativos y promover la evaluación como instrumento para generación de conocimiento (UNEG/FAO 2012, pág. x).

## **Grupo 2: Calidad en las evaluaciones a partir de la institucionalización de la función de evaluación, y de las políticas y prácticas evaluativas (PP)**

En este segundo grupo nos encontramos a tres entidades, FMAM, OIOS y ONU Mujeres que han logrado resultados positivos en cuanto a la calidad de sus evaluaciones. Dentro de este segundo grupo, hay una diferencia fundamental: sus niveles en cuanto a la cultura de aprendizaje. En este sentido, OIOS y FMAM presentan niveles algo menos consolidados que ONU Mujeres y, sin embargo, esto no ha hecho mermar sus niveles de calidad. A continuación presentamos las características de estas tres agencias.

El grado de institucionalización de la función de evaluación se ha medido a través del grado de independencia estructural y de recursos, tanto financieros como humanos que despliegan las agencias. Comenzando por los niveles de independencia estructural, observamos que las tres entidades gozan de una independencia estructural muy consolidada<sup>90</sup>.

Estas organizaciones comparten altos grados de independencia en sus respectivas funciones de evaluación aunque por diferentes razones. En el caso de la OIOS, la convivencia de la función de evaluación dentro de la misma unidad con otras funciones consideradas de 'control', como por ejemplo auditoría, investigaciones e inspecciones, otorga un carácter independiente a la práctica evaluativa, ya que en el propósito de rendición de cuentas (*accountability*) adquiere un papel más relevante que el propósito del aprendizaje. En opinión del panel de revisión por pares de OIOS *"la percepción de*

---

<sup>90</sup>En el caso de FMAM, esta independencia estructural todavía no estaba legalmente formalizada en la segunda revisión por pares en Mayo de 2014. Por su parte, los informes realizados por la Sección de Evaluación e Inspecciones de OIOS son reportados al Quinto Comité de la Asamblea General a través del Secretario General. ONU Mujeres tiene canal de comunicación directo con la Junta Ejecutiva.

*la evaluación en OIOS como un instrumento primordialmente de supervisión y rendición de cuentas ha dificultado la dimensión de la evaluación que la identifica con una herramienta de aprendizaje y como un vehículo de cambio*”, (UNEG/OIOS 2009, pág. 11). En el caso del FMAM, nos encontramos que sus niveles de independencia incluyen no solo un canal de comunicación directo con el Consejo (órgano de gobierno), sino que también conllevan la elección directa del director/a de la unidad de evaluación por los miembros del Consejo. Es decir, la independencia es completa. ONU Mujeres, por su parte, a pesar de que a nivel interno la división de evaluación reporte a la Dirección General, el canal de comunicación con la Junta Ejecutiva es directo.

La independencia estructural que los tres organismos despliegan, viene acompañada a su vez de una elevada independencia financiera al disponer de capacidad plena de planear y gestionar sus recursos. Los presupuestos para hacer frente al plan de evaluación son accesibles y transparentes, y son aprobados por sus respectivos órganos de gobierno (El Consejo en el caso del FMAM, la Junta Ejecutiva para ONU Mujeres y el Quinto Comité de la Asamblea General en el caso de OIOS). Los paneles de revisión por pares coinciden en considerar que la asignación presupuestaria a evaluación es suficiente en ambos casos para cumplir de manera satisfactoria con el plan de trabajo, aunque también previene que es necesario un ejercicio más enfocado en la priorización para poder mantener los estándares de calidad existentes<sup>91</sup>.

Este alto grado de institucionalización también se traduce en una gestión efectiva de los recursos humanos. Así, el incremento de la carga de trabajo de las unidades de evaluación ha ido emparejado con un incremento del presupuesto regular con la consecuente contratación de personal adicional. En el caso de OIOS se tiende a prescindir de la contratación de personal externo<sup>92</sup> y el FMAM limita la presencia de personal externo a salvaguardar los requerimientos de conocimientos técnicos específicos que, en momentos determinados, no existan en la unidad de evaluación.

Sin embargo, los paneles de revisión por pares destacan que, en el caso de OIOS y FMAM, los altos niveles de independencia estructural, financiera y organizativa que presentan las dos organizaciones pueden haber llevado a sus unidades de evaluación a una situación de aislamiento y distancia con las entidades implementadoras<sup>93</sup> (UNEG/OIOS 2009, págs. 31, 39; UNEG/GEF 2009, págs. 45-46<sup>94</sup>). Esto coincide con la relativa poca importancia otorgada al factor aprendizaje que se observa en las dos

---

<sup>91</sup>Así, a pesar de que el presupuesto sea transparente y haya aumentado progresivamente, se considera que el mandato de OIOS de cubrir ‘todas las actividades implementadas bajo la autoridad de la Secretaría’ es muy amplio, y puede ser insuficiente (UNEG/OIOS 2009, pág. 30).

<sup>92</sup>En ocasiones, esta tendencia puede haber ocasionado ciertas deficiencias técnicas en la composición de los equipos de evaluación (UNEG/OIOS 2009, pág. 44)

<sup>93</sup>Es importante destacar la particularidad de estas dos entidades con respecto a las otras entidades objeto de estudio. El FMAM no es una agencia ejecutora de fondos, es un mecanismo financiero y como tal las responsabilidades de la Secretaría de FMAM son principalmente de asesoría, supervisión y enlace con las agencias implementadoras de fondos FMAM. Por su parte, OIOS tampoco ejecuta fondos, ni proporciona financiación, sino que es un mecanismo de supervisión.

<sup>94</sup>Aunque se observa cierta mejoría en el segundo proceso de revisión por pares, el panel apunta a que la participación de las unidades técnicas de la Secretaría del GEF y las agencias implementadoras de FMAM todavía es limitada (UNEG/GEF 2014, pág. 10).

organizaciones frente a la rendición de cuentas (en contraposición con ONU Mujeres). A continuación se presentan algunos detalles:

En ambos casos (FMAM y OIOS), los paneles de revisión por pares concluyen en sus observaciones que en las dos agencias no se observan claros signos de aplicación de una estrategia de disseminación de la información evaluativa de manera consistente. El FMAM tiene un repositorio electrónico de las evaluaciones que es accesible públicamente pero, más allá de esta herramienta de consulta, hay una muy limitada evidencia de otras actividades de disseminación activa. OIOS, por su parte, efectúa una disseminación esporádica que incluye la publicación restringida de los informes que se presentan a los órganos de gobierno los cuales representan solo una parte de las evaluaciones que se realizan.

Los bajos niveles de participación, implicación y consulta en las evaluaciones son señalados por los paneles de revisión por pares como los aspectos más críticos de la función de evaluación en ambas organizaciones. A nivel de contrapartes nacionales, se observa gran nivel de desconocimiento sobre las evaluaciones. A nivel organizacional, se observa una participación inconsistente, enfocada principalmente en algunas etapas (como por ejemplo en los ToRs en el caso de OIOS) y una consideración generalizada de que las evaluaciones se llevan a cabo con un enfoque de arriba abajo (*top down*). Este hecho también se refleja en sus procedimientos operacionales: por ejemplo, en el caso de FMAM, el tiempo para proporcionar comentarios a los informes de evaluación es considerado insuficiente, lo mismo que para presentar el plan de acción gerencial (UNEG/GEF 2009 págs. 10, 12, 15, 41, 59, 69).

Este distanciamiento de la función de evaluación se traduce en una percepción algo negativa y, en muchos casos, confusa sobre lo que la evaluación puede aportar a la organización. En el caso de OIOS, el personal entrevistado y los Miembros del Quinto Comité encuentran dificultades para distinguir evaluación de otras funciones propias de la supervisión como auditoría o investigaciones (UNEG/OIOS 2009, pág. 42). También en el caso del FMAM, el panel de revisión por pares no observa un empoderamiento derivado de los resultados de la evaluación por parte de la Secretaría. Al contrario, incluso se señala cierto nivel de competitividad entre la Secretaría y la Oficina de Evaluación por lograr la atención y prioridad de los órganos de gobierno (UNEG/GEF 2009; págs. 13, 69, UNEF/GEF 2014, pág.17).

Los resultados obtenidos en el análisis de revisión por pares permiten evidenciar que los altos niveles de institucionalización y de esfuerzo por cultivar una cultura de independencia son más propios en organizaciones donde el público receptor (o usuarios intencionados) de las evaluaciones está más claramente vinculado al nivel político y, por tanto, a sus órganos de gobierno, lo que en consecuencia provoca que se priorice el propósito de la rendición de cuentas frente al de aprendizaje organizacional. Este hecho ha podido tener importantes consecuencias en la desvinculación y distancia de las unidades de evaluación con el personal de la organización y en general, con otros actores (donantes, usuarios finales, contrapartes nacionales, agencias/entidades implementadoras). Sin embargo, no ha disminuido el

resultado de calidad de las evaluaciones. De hecho, estos resultados confirman que también es posible llegar a altos estándares de calidad, en situaciones donde la participación y consulta sean limitadas.

Así, los paneles de revisión por pares observan claros signos de calidad en los informes de evaluación de las dos agencias, principalmente en el FMAM, donde los informes incluyen sólida información técnica, hallazgos basados en evidencias rigurosas y análisis profundos y detallados. El rigor metodológico está presente en las evaluaciones analizadas. Las entrevistas realizadas en los procesos de revisión por pares confirman la apreciación de las personas entrevistadas por la calidad de los informes de las dos agencias, aunque también se observa que la valoración más positiva proviene de los receptores esperados, en este caso, los órganos de gobierno (UNEG/GEF 2009, págs. 12, 41, 46, 57; UNEG/OIOS 2009, pág. 47).

Por lo tanto se puede afirmar que en situaciones de alta institucionalización de la función de evaluación, incluso en aquellos casos donde ésta haya podido resultar en cierto aislamiento, la calidad es posible debido a un factor fundamental: la presencia de una sólida aplicación de políticas y prácticas evaluativas.

Si analizamos los tres casos de FMAM, OIOS y ONU Mujeres vemos que los tres coinciden en contar con políticas sólidas de evaluación (en el caso del FMAM aprobada por el Consejo en 2006 y actualizada en 2010, en el caso de OIOS la PPBME aprobada en 2008 por la Asamblea General, y en el caso de ONU Mujeres aprobada en 2012 por la Junta Ejecutiva), así como con planes de evaluación elaborados en base a criterios multidisciplinares preestablecidos, y un sistema de seguimiento de las recomendaciones, cuyo análisis es presentado a los órganos de gobierno<sup>95</sup>. Por lo tanto, podemos intuir que, efectivamente, las políticas y prácticas evaluativas son una condición necesaria, que hace que el resto de factores (en este caso la institucionalización) sean efectivos.

La importancia de las políticas y prácticas evaluativas en la calidad de la evaluación también se ha visto corroborada en otros casos, donde se han llevado a cabo diferentes procesos de revisión por pares, y se ha podido ver una evolución de las políticas y prácticas evaluativas a través del tiempo. Por ejemplo, UNICEF y PMA son dos agencias que efectuaron sus procesos de revisión por pares en 2006, y han llevado procesos de revisión de sus funciones de evaluación posteriormente. En este sentido, en ambas agencias, la aprobación de sus Políticas de Evaluación por sus respectivas Juntas Ejecutivas en 2008 y la consecuente puesta en marcha de medidas para la efectiva aplicación de la misma (criterios claros y transparentes en la selección de

---

<sup>95</sup> El sistema '*Management Action Record* (MAR)' es una base de datos creada por el FMAM que incluye la adopción de decisiones por parte del Consejo y su relación con las recomendaciones de las evaluaciones. El primer informe de revisión por pares señala que podría mejorarse la parte cualitativa del sistema de seguimiento para darle una aplicación más cualitativa/sustantiva y menos estadística. (UNEG/GEF 2009, págs. 12, 58, 66). El segundo informe, apunta a que su aplicación está muy orientada en la rendición de cuentas, y no tanto al aprendizaje. (UNEG/GEF 2014, pág. 17). Estas conclusiones también coinciden con el panel de revisión por pares de OIOS. El sistema de seguimiento a las recomendaciones de OIOS existe y se aplica (*Issue Track*), pero su aplicación esta principalmente orientada a la rendición de cuentas, no al aprendizaje (UNEG/OIOS 2009, pág. 37)

evaluaciones, canales de reportes directos con la Junta Ejecutiva, sistema de seguimiento a las recomendaciones) ha supuesto un antes y un después con respecto a la mejora de la calidad de sus evaluaciones y la puesta en marcha de controles más robustos para monitorear la misma (Universalía 2013, pág. iv; UNEG/WFP 2014, pág. 11)<sup>96</sup>.

---

<sup>96</sup>El Sistema Geros (*Global Evaluation Reports Oversight System*), es un sistema que analiza la calidad de los informes finales de evaluación, en base a unos estándares de calidad previamente establecidos, y se realiza por una compañía externa e independiente. A través de este sistema también se producen meta evaluaciones, que son disseminadas en la web y discutidas en la Junta Ejecutiva.

## 5.2. El uso sostenible de la evaluación

El uso sostenible de la evaluación constituye el segundo resultado de la capacidad evaluativa en nuestro modelo de análisis. Tal y como se ha procedido con el resultado de calidad, primero se analizará la existencia de condiciones de necesidad (a través de índices de consistencia y de los diagramas XY), y después se realizará el análisis de las condiciones de suficiencia, a partir de los resultados procedentes de la tabla de la verdad. Estos resultados nos permitirán discernir posibles combinaciones causales (o caminos), considerados como suficientes para poder lograr un uso sostenible de la evaluación en las organizaciones.

En relación al uso sostenible, es importante precisar que el tipo de uso principalmente considerado en nuestro modelo es el uso de tipo instrumental, aquel que es posible verificar a través de la generación de acciones y cambios tangibles. Para su medición, se han propuesto tres indicadores fundamentales: 1) Generación de cambios, 2) Implementación de recomendaciones, y 3) Relevancia de las evaluaciones en los foros de toma de decisiones. Tal y como se ha procedido para el análisis de la calidad, la información se ha obtenido de los informes de revisión por pares efectuados para las agencias de nuestra muestra.

### 5.2.1 Análisis de las condiciones necesarias

De la misma manera que con el resultado de la calidad, a continuación se procede a analizar la existencia de condiciones necesarias para lograr un uso sostenible de la evaluación (Tabla 5.4).

**Tabla 5.4. Análisis de condiciones necesarias para el resultado ‘uso de la evaluación’.**

Condición	Consistencia	Cobertura
CA	1,000	1,000
ca	0,722	0,620
INS	0,785	0,731
ins	0,859	0,787
PP	0,857	0,800
pp	0,792	0,724

Nota: La tabla 5.4 describe los índices de cobertura y consistencia de las condiciones objeto de estudio (CA, INS, y PP), y también de la ausencia de dichas condiciones (ca, ins y pp).

En este caso, observamos que la capacidad de aprendizaje (CA) se constituye en condición necesaria para lograr un uso sostenible de la evaluación. El índice de consistencia máximo implica que en todos los casos donde la condición ha estado presente (es decir, todas las agencias con capacidad de aprendizaje), los resultados son positivos en cuanto al uso de la evaluación.

La Tabla 5.4 también nos muestra un índice de cobertura total en dicha condición. Como se ha mencionado anteriormente, la cobertura refleja el grado en que la condición en cuestión está presente o cubre los casos donde el resultado objeto de

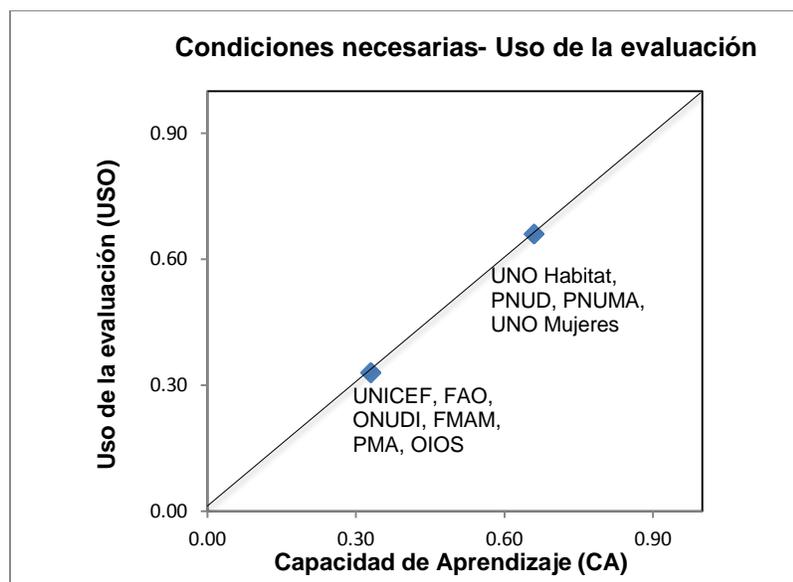
estudio se produce. Por tanto, la cobertura mide la relevancia explicativa de la condición. Así, una cobertura de 1 implica que, en todos los casos donde el resultado se produce, la condición está presente. Dicho de otro modo, no existe en nuestra muestra un solo caso donde se consigan resultados positivos en cuanto al uso de la evaluación y que no sea debido a la existencia de capacidad de aprendizaje en la organización.

También se desprende otro dato interesante a priori y que será analizado más en detalle gráficamente. En efecto, vemos que la ausencia de institucionalización de la función de evaluación (ins) puede ser incluso más determinante en el uso de la evaluación que su presencia (INS) (con índices de consistencia respectivamente de 0,859 y 0,785).

En cuanto a la consistencia de las condiciones, la capacidad de aprendizaje (CA) es seguida por dos condiciones que, a nivel individual, son prácticamente igual de relevantes, las políticas y prácticas evaluativas (PP) y la ausencia de institucionalización (ins). En el apartado dedicado al análisis de suficiencia se analizarán las combinaciones entre estas tres condiciones.

Tal y como procedimos con el análisis de la calidad, el análisis de la necesidad de las condiciones también se ha complementado gráficamente. Los gráficos 5.6, 5.7 y 5.8 muestran la relación entre cada una de las condiciones y el resultado 'uso de la evaluación'.

**Gráfico 5.6. Capacidad de Aprendizaje y Uso de la Evaluación**



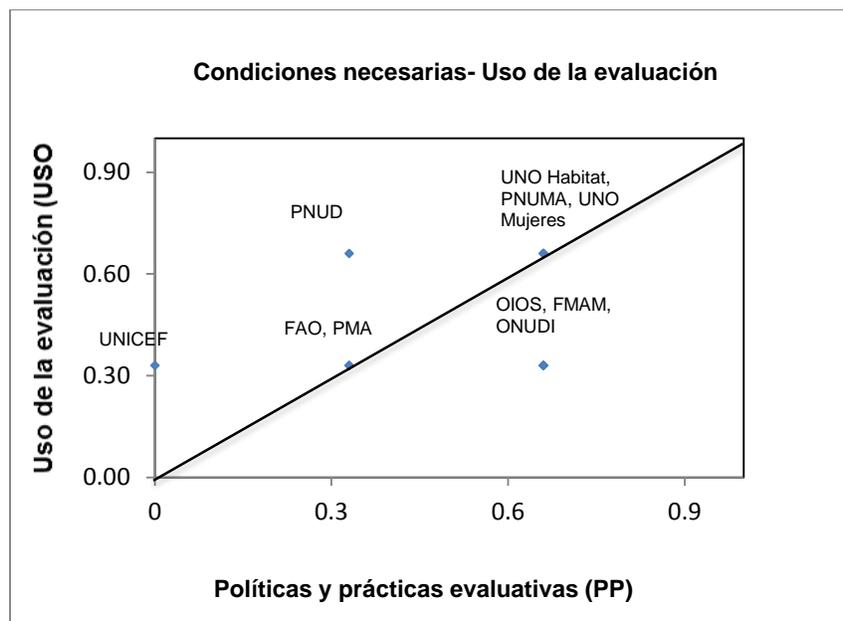
El gráfico 5.6 nos permite ver la estrecha relación existente entre la capacidad de aprendizaje y el logro de resultados en utilización, puesto que todas las agencias

obtienen resultados que se encuentran sobre la línea divisoria y, por tanto, tienen el mismo nivel de pertenencia tanto de la condición al subconjunto, como al resultado.

En total, existen cuatro agencias que presentan altos niveles en cuanto a su capacidad de aprendizaje y, al mismo tiempo, presentan altos niveles de uso de la evaluación: UNO Habitat, PNUD, PNUMA y UNO Mujeres. A pesar de ciertas diferencias en cuanto a sus estrategias de disseminación y sus procesos de consulta, así como en relación con la motivación y predisposición del personal en cuanto a práctica evaluativa, las cuatro agencias coinciden en haber obtenido resultados positivos en cuanto al uso de la evaluación (UNEG/Habitat 2012, págs. v, vi, 31; UNEG/UNDP 2006, págs., 11, 19, 45; UNEG/UNEP 2012, págs. 16, 30-31, UNEG/UN Women, págs.12, 22, 23).

Por su parte, un total de seis agencias obtienen bajos niveles, tanto en capacidad de aprendizaje como en la utilización de la evaluación. Se trata de UNICEF, FAO, ONUDI, FMAM, PMA y OIOS. Principalmente coinciden en una participación inconsistente del personal involucrado y las contrapartes nacionales en los procesos de evaluación, limitaciones en la aplicación de una estrategia de disseminación adecuada y una cultura evaluativa todavía en proceso de consolidación en estas organizaciones.

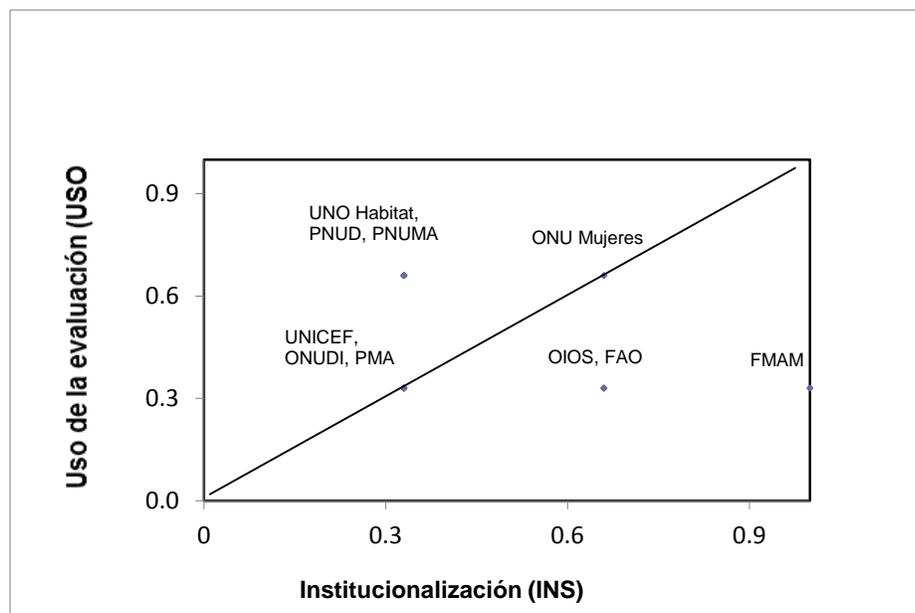
**Gráfico 5.7. Políticas/Prácticas evaluativas y Uso de la Evaluación**



El análisis del Gráfico 5.7 corrobora los datos de consistencia de las condiciones de necesidad para el resultado relativo al uso de la evaluación. Así, analizando en concreto la condición de políticas y prácticas evaluativas (PP), vemos que un total de cinco agencias se posicionan sobre la línea divisoria, presentando los mismos resultados con respecto a la condición que al resultado. Tres de ellas consiguen resultados positivos en ambos (0,66 para UNO Habitat, PNUMA y UNO Mujeres), y dos (FAO, y PMA), presentan bajos resultados tanto en la condición como en el resultado.

Por su parte, el PNUD es una agencia que, a pesar de que, cuando se realizó el proceso de revisión por pares, sus políticas y prácticas evaluativas no estaban aún plenamente establecidas, sin embargo los resultados en cuanto al uso son positivos, lo que descarta a las políticas y prácticas evaluativas como condición necesaria. Como contraparte, nos encontramos con agencias como OIOS, FMAM y ONUDI donde, a pesar de contar con sólidas políticas y prácticas evaluativas, no se perciben resultados tan positivos con respecto al uso de la evaluación, lo que descarta la condición como suficiente.

**Grafico 5.8. Institucionalización y Uso de la Evaluación**



Finalmente, el análisis gráfico de la condición de Institucionalización de la función de evaluación, también permite comprobar que no se trata de una condición necesaria o suficiente por sí misma para el logro de resultados en cuanto a utilización. Cuatro de las 10 agencias (ONU Mujeres con resultados positivos en ambos y UNICEF, ONUDI y PMA con resultados negativos) se encuentran sobre la línea divisoria. Sin embargo, tres agencias consiguen resultados positivos en cuanto al uso a pesar de contar con estructuras de evaluación poco consolidadas en sus organizaciones (ONU Habitat, PNUD, PNUMA), por lo que no se trata de una condición necesaria. Por su parte, la existencia de tres agencias (OIOS, FAO y FMAM) con altos niveles de institucionalización en la organización, pero que no obtienen resultados tan positivos en cuanto al uso, inducen a considerar que la institucionalización en sí no es suficiente para lograr un uso sostenible de la evaluación.

## 5.2.2 Análisis de las condiciones suficientes

El análisis de las condiciones suficientes para que se produzca el resultado ‘uso sostenible de la evaluación’, se realiza igualmente mediante el algoritmo de la tabla de la verdad de Quine McClusky (Schneider y Wagemann 2007). A través de la tabla de la verdad, vemos que una combinación es suficiente para que se produzca el resultado cuando la pertenencia de cada miembro de la muestra (organismo) al conjunto conformado por la combinación de condiciones es igual o menor que su pertenencia al conjunto representado por el resultado. A continuación, la Tabla 5.5 presenta la tabla de la verdad para el resultado ‘uso sostenible de la evaluación’.

**Tabla 5.5. Tabla de la verdad sobre el resultado ‘Uso de la Evaluación’.**

CA	INS	PP	Agencias	Casos	USO	Consistencia
1	0	0	PNUD	1	1	1,000
1	1	1	UNO Mujeres	1	1	1,000
1	0	1	UNO Habitat, PNUMA	2	1	1,000
0	1	0	FAO	1	0	0,904
0	0	1	UNIDO	1	0	0,887
0	1	1	OIOS, FMAM	2	0	0,818
0	0	0	UNICEF, PMA	2	0	0,809

Nota: El umbral de consistencia para las condiciones suficientes se ha establecido en 0.95

Asimismo, la Tabla 5.6 muestra los resultados del análisis de las condiciones suficientes del resultado ‘uso sostenible de la evaluación’.

**Tabla 5.6. Análisis de condiciones suficientes en ‘Uso de la Evaluación’.**

Solución	Compleja		Parsimoniosa
	CA*PP	CA*ins	CA
Agencias	UNO Habitat PNUMA UNO Mujeres	UNO Habitat PNUMA PNUD	UNO Habitat PNUMA UNO Mujeres PNUD
Consistencia	1,000	1,000	1,000
Cobertura total	0,857	0,859	1,000
Cobertura exclusiva	0,140	0,142	1,000
Consistencia de la solución	1,000		1,000
Cobertura de la solución	1,000		1,000

La tabla de la verdad nos ofrece dos soluciones (la compleja y la parsimoniosa) y ambas obtienen niveles máximos de consistencia y cobertura (1,000 respectivamente). Esto implica que los caminos ofrecidos para conseguir un uso sostenible de la evaluación no presentan excepciones (casos donde ante la combinación de condiciones no se produzca el resultado), y su carácter explicativo es total, ya que siempre que se produce el resultado se produce la solución propuesta.

Para la solución compleja, es decir, en el supuesto en el que el contra fáctico de nuestro estudio (presencia de capacidad de aprendizaje, presencia de institucionalización, y ausencia de políticas/prácticas evaluativas) no produjera resultados en cuanto al uso de la evaluación, la tabla de la verdad identifica dos 'camino', o condiciones suficientes que llevan a altos niveles de uso de la evaluación.

Los dos caminos considerados como suficientes en la solución compleja son:

Condición suficiente 1: Alta capacidad de aprendizaje (CA) combinada con alto nivel de políticas y prácticas evaluativas (PP)

Condición suficiente 2: Alta Capacidad de aprendizaje (CA) combinada con bajo nivel de institucionalización (ins)

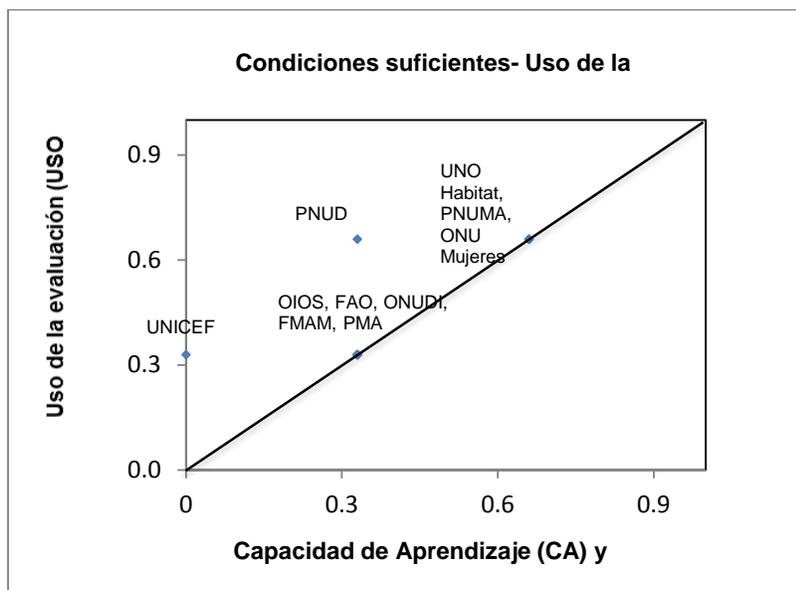
Estos resultados que ofrece la tabla de la verdad confirman la condición de capacidad de aprendizaje como condición necesaria, puesto que se encuentra presente en las dos soluciones.

Por su parte, en la solución parsimoniosa, que implica que en el contra fáctico se produzca un resultado positivo con respecto al uso de la evaluación, la tabla de la verdad identifica la capacidad de aprendizaje, como una condición necesaria y suficiente para el logro de resultados positivos de uso.

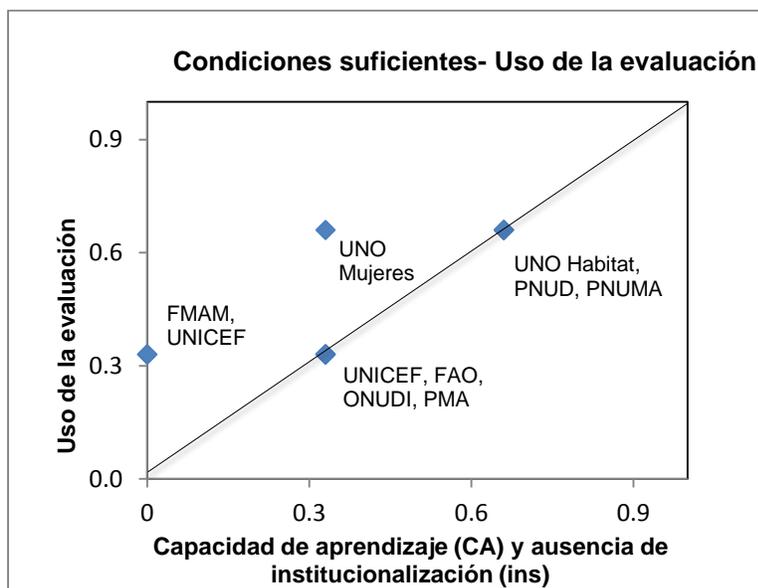
En base a estos resultados, la consideración del carácter del contra fáctico es importante a la hora de llegar a conclusiones. La solución parsimoniosa considera irrelevante la presencia o ausencia de las otras dos condiciones (institucionalización y políticas/prácticas evaluativas) para el logro de resultados. Sin embargo, la solución compleja propone una combinación de condiciones donde la capacidad de aprendizaje es fundamental, pero no suficiente para el logro de resultados positivos en uso y se debe complementar con políticas y prácticas evaluativas, o bien con bajos niveles de institucionalización de la función de evaluación. De cara a enriquecer las conclusiones del análisis, a continuación se procede a analizar los dos caminos ofrecidos por la solución compleja.

Para ello, primeramente, procedemos a presentar los resultados gráficamente (Gráficos 5.9 y 5.10). Recordamos que, para que una condición sea suficiente, todos los casos deben estar situados alrededor o sobre la línea divisoria (Ragin 2000, pág. 236).

**Gráfico 5.9. El uso de la evaluación desde la capacidad de aprendizaje y las políticas/prácticas evaluativas**



**Gráfico 5.10. El uso de la evaluación desde la capacidad de aprendizaje y en ausencia de institucionalización**



Del análisis de suficiencia en el uso de las evaluaciones, se desprenden por tanto las siguientes conclusiones:

La capacidad de aprendizaje se confirma como condición necesaria, y en función de la consideración que se haga del contra factico, podría resultar suficiente. Por tanto, la capacidad de aprendizaje forma parte de todas las posibles soluciones que acercan a las organizaciones al uso sostenible de las evaluaciones.

Las dos soluciones propuestas en la solución compleja presentan altos niveles de duplicidad. Es decir, un total de 4 agencias coinciden en obtener resultados satisfactorios en cuanto al uso de la evaluación (UNO Habitat, PNUMA, UNO Mujeres y PNUD). Sin embargo, UNO Habitat y PNUMA forman parte de las dos soluciones propuestas, puesto que ambas coinciden en su combinación de condiciones (CA\*PP\*ins), mientras que UNO Mujeres y PNUD están presentes en solo una de las dos soluciones, ya que tienen una combinación única diferenciada (UNO Mujeres presenta la combinación CA\*PP\*INS, y PNUD presenta la combinación CA\*ins\*pp).

A continuación se explican más en detalle las características de cada uno de los dos caminos ofrecidos para llegar al logro de resultados positivos en utilización desde la solución compleja:

### **Grupo 1: El camino hacia la utilización desde el aprendizaje (CA) y las políticas y prácticas evaluativas (PP)**

En este grupo se encuentran tres agencias ya analizadas en detalle en el apartado anterior (UNO Habitat, PNUMA y UNO Mujeres), puesto que este camino (que combina aprendizaje con políticas y prácticas evaluativas), también coincide con una de las dos soluciones propuestas para el logro de la calidad.

Las tres agencias coinciden por haber realizado una clara apuesta por los procesos participativos y el aprendizaje como propósito evaluativo (UNEG/UN Habitat 2012, págs. 13, 25 y 29; UNEG/UNEP 2012, pag.19, 23, 25 y 30; UNEG/UN Women 2014, págs. 3, 12). Por ello, aunque todavía se puede incrementar la participación e implicación de los países (contrapartes nacionales) y de las oficinas a nivel de país, tanto en la planificación como en la discusión de los hallazgos, los esfuerzos destinados por estas agencias a promocionar la evaluación como herramienta de aprendizaje son apreciados por su propio personal. Asimismo, estos esfuerzos han sido combinados con el establecimiento de un marco normativo que responde a los estándares UNEG en cuanto al desarrollo de su política de evaluación, plan de evaluación y sistema de seguimiento de las recomendaciones (UNEG/UN Habitat 2012, págs., 11, 14, 16 y 19; UNEG/UNEP 2012, págs., 16, 19, 30-33; UNEG/UN Women 2014, págs. 7, 17, 23).

En cuanto al uso de la evaluación se puede observar lo siguiente:

- ✓ Especialmente en el caso de UNO Habitat y PNUMA, y en menor medida en UNO Mujeres (puesto que el establecimiento de la función de evaluación es más reciente), existe una clara percepción de que las evaluaciones tienen influencia en la toma de decisiones y que los hallazgos y las recomendaciones, a menudo,

son consideradas en la planificación del nuevo ciclo programático. Sin embargo, las evaluaciones de carácter estratégico reciben menos atención lo que provoca que los cambios generados a partir de la práctica evaluativa sean de tipo operativo, en vez de ubicarse a niveles de políticas o estrategias. En general, se considera que la evaluación ha sido influyente en la promoción de cambios a nivel organizativo y en la validación o ajustes de orientaciones operativas (programas y proyectos) (UNEG/Habitat 2012, págs. v, vi; UNEG/UNEP 2012, pág. 30)

- ✓ Las tres agencias coinciden en el alto porcentaje de recomendaciones implementadas, en los casos donde existe información al respecto, el porcentaje se acerca al 70%. En el caso de UNO Mujeres, esta información es públicamente accesible a través del sistema GATE. En el caso de PNUMA, la gerencia de la organización incluye como medida de valoración del rendimiento de los/as directores/as de los programas las estadísticas relativas al estado de las recomendaciones.
- ✓ Se observa que las evaluaciones están presentes en los procesos de toma de decisiones especialmente en el caso de PNUD y UNO Habitat<sup>97</sup>. La presentación y el análisis de las evaluaciones tiene lugar a nivel de los órganos de gobierno (El Comité de Representantes Permanentes en el caso de UNO Habitat y la Junta Ejecutiva en el caso de PNUD) e internamente. El papel relevante de la evaluación en dichos foros es compartido por los órganos de gobierno y la alta gerencia.
- ✓ Es significativo el nivel de consistencia que existe entre las 2 condiciones y el resultado (máximo). Así, en todos los casos donde existe CA y PP, existen resultados positivos relativos al uso de la evaluación. Estos datos reafirman las teorías existentes que relacionan el aprendizaje y el marco normativo con la utilidad.

A partir de los resultados expuestos emerge el siguiente interrogante: ¿qué grado de relevancia ocupa la institucionalización de la evaluación en esta solución? En el caso de UNO Habitat y PNUMA, los niveles de institucionalización son débiles, al contrario de la situación en UNO Mujeres. Las dos principales carencias de las dos primeras (UNO Habitat y PNUMA) con respecto a UNO Mujeres en el ámbito institucional son: i) ausencia de canal directo de comunicación entre la unidad de evaluación y los órganos de gobierno y ii) el presupuesto se considera insuficiente para hacer frente al plan de evaluación y la unidad de evaluación no tiene pleno control sobre el mismo para la toma de decisiones. Estos datos permiten indicar que, a priori, la institucionalización, o

---

<sup>97</sup>En el caso de PNUMA, se percibe que todavía se puede mejorar los niveles de comunicación entre la Oficina de Evaluación y los órganos de gobierno. De hecho, la información sobre evaluación es reportada al Director Ejecutivo a través del informe de síntesis bianual (UNEG/UNEP 2012, pág. 20).

su ausencia, puede no ser un factor tan determinante del uso si la capacidad de aprendizaje y el marco normativo presentan niveles sólidos.

## **Grupo 2: El camino hacia la utilización desde el aprendizaje y en ausencia de institucionalización.**

Se incluyen en este grupo dos de las tres agencias mencionadas en el grupo anterior (UNO Habitat y PNUMA) además del PNUD. Las tres coinciden en lograr resultados positivos en cuanto a la utilidad de sus evaluaciones.

PNUD coincide con UNO Habitat y PNUMA en apostar por la promoción de la evaluación como herramienta para el aprendizaje y la disseminación. Esta estrategia incluye el acceso público a las evaluaciones centralizadas y descentralizadas, y la disseminación a través de redes<sup>98</sup>. También se realizan síntesis de informes de evaluación, bien como herramienta promocional y de comunicación (llamados *'digest'*), o como las síntesis que se incluyen en los informes anuales a la Junta Ejecutiva<sup>99</sup>.

Estos esfuerzos de comunicación fueron valorados positivamente por el panel de revisión por pares de PNUD, especialmente sus efectos positivos en la creación de un ambiente de sintonía entre la alta gerencia, el personal y los órganos de gobierno en relación a los propósitos de la evaluación. En general, la revisión por pares destaca una significativa apreciación entre los miembros de la Junta Ejecutiva de la información proveniente de la Oficina de Evaluación<sup>100</sup>. La creciente expectación de la Junta Ejecutiva por las evaluaciones corporativas (es decir, las que son efectuadas a nivel central y no están descentralizadas a las oficinas de campo) ha sido algo ya señalado en la revisión por pares efectuada en 2006, y reiterado en la efectuada en 2013 (UNEG/UNDP 2006, pág. 45, UNEG/UNDP 2013, pág.9).

Entre el personal de PNUD, existe así la percepción que el propósito del aprendizaje organizacional prima sobre el de la rendición de cuentas (UNEG/UNDP 2006, pág. 32). Igualmente, existe una valoración positiva sobre el valor añadido que la evaluación puede proporcionar a la formulación de nuevos programas y al aprendizaje dentro del propio organismo.

PNUD también presenta similares limitaciones a las encontradas en UNO Habitat y PNUMA con respecto al nivel de institucionalización de la función de evaluación que se

---

<sup>98</sup> Redes como 'UNDP Network' y ' Organization Evaluation Forum'

<sup>99</sup>Estos esfuerzos de disseminación también han sido analizados en detalle en los ejercicios de revisión por pares y análisis de la función de evaluación del PNUD de 2013 y 2014 (UNEG/UNDP 2013, Baasterl 2014). Como conclusión, teniendo en cuenta los avances de la Oficina de Evaluación en comunicación, se señala, que PNUD debe reorientar su política de conocimiento hacia un enfoque más bidireccional para lograr un mayor impacto: *"Mas allá de la gestión del conocimiento 'knowledge management', la función de evaluación debe dirigirse hacia una mayor interacción 'knowledge sharing', desde un enfoque bidireccional e ilimitado del proceso de comunicación"* (UNEG/UNDP 2013, pág. 17).

<sup>100</sup>Los informes de evaluación son accesibles a través de la red. Sin embargo, las evaluaciones presentadas a la Junta Ejecutiva son las corporativas, no las que se realizan de manera descentralizada.

reflejan en una ausencia de canal directo de comunicación entre la unidad de evaluación y la Junta Ejecutiva, así como en recursos insuficientes para la función evaluativa. En particular, con respecto al canal de reporte, se puede afirmar, a pesar de contar con independencia funcional, no hay canal de comunicación directa con la Junta Ejecutiva puesto que la Oficina de Evaluación reporta a la Junta Ejecutiva a través del Administrador/a<sup>101</sup>, persona que presenta dos tipos de informes a la Junta Ejecutiva sobre evaluación: el informe anual sobre evaluación y el informe bianual sobre la efectividad de la evaluación. En estos informes se incluyen síntesis sobre evaluaciones realizadas, datos sobre seguimiento de las evaluaciones y sus recomendaciones, así como información sobre el plan de evaluación y los resultados de las evaluaciones corporativas.

En cuanto a la insuficiencia de recursos, la revisión por pares de 2006 de PNUD concluye que la asignación de recursos proviene de fuentes diferentes y no existe un consolidado de las mismas. La Junta Ejecutiva no conoce la magnitud del plan de evaluación (puesto que solo se incluyen en el mismo las evaluaciones corporativas). Esto implica que las oficinas de campo y las oficinas regionales son las encargadas de asignar presupuesto a las evaluaciones descentralizadas. En consecuencia, las evaluaciones descentralizadas (evaluaciones de país y de proyecto), al no formar parte del presupuesto regular, no tienen un presupuesto garantizado. La Política de Evaluación aprobada en el 2006 estableció la obligatoriedad de incluir una línea presupuestaria para evaluación en el presupuesto de todos los proyectos. Aun así, esta medida no ha resuelto las limitaciones presupuestarias todavía existentes y, por tanto, se identifica un nivel insuficiente de recursos y capacidades en la evaluación descentralizada del PNUD (Baasterl 2014, UNEG/UNDP 2013).

A pesar de las limitaciones anteriores, ya en la revisión por pares efectuada en el 2005, el panel de revisión por pares de PNUD sí que advierte de la presencia de indicadores precisos de intencionalidad en cuanto al uso de la evaluación (UNEG/UNDP 2006, pág. 45). El interés de la Junta Ejecutiva por la evaluación está favoreciendo positivamente su relevancia en la toma de decisiones. Las evaluaciones, en particular las corporativas, son discutidas en foros de gerencia y tienen relevancia en la toma de decisiones. En concreto, las deliberaciones de la Junta Ejecutiva con respecto a las evaluaciones efectuadas sobre el Segundo Marco Global de Cooperación y el Marco de Cooperación Regional de los Estados Árabes, estuvieron directamente ligadas a las decisiones de los sucesivos marcos de programa (UNEG/UNDP 2006). Esto se corrobora con los hallazgos del análisis de la función de evaluación efectuado en 2014, donde se documenta la forma en que las evaluaciones corporativas han influenciado cambios y decisiones en el Plan Estratégico de 2014-2017 (Baasterl 2014, pág. 21).

---

<sup>101</sup>La ambigüedad en las responsabilidades del Director/a de la Oficina de Evaluación y su falta de reporte de comunicación directo con la Junta Ejecutiva ha sido mencionada en los informes de revisión por pares como una limitación que puede afectar la independencia operacional de la Oficina de Evaluación (UNEG/UNDP 2006 pág. 9, 12; Baasterl 2014, pág., ix, 71).

Por lo tanto, los resultados de este segundo grupo evidencian que la capacidad de aprendizaje combinada con bajos niveles de institucionalización también puede derivar en evaluaciones que tengan capacidad de influencia y capacidad de generar cambios desde diferentes ángulos:

- ✓ Un nivel de independencia y distanciamiento no tan marcado entre unidades técnicas y unidad de evaluación puede ayudar a crear una percepción de mayor cercanía de la unidad de evaluación con el personal de la organización, lo que se traduce en un mayor énfasis en el aprendizaje como propósito de la evaluación sobre la rendición de cuentas y una mayor apreciación del potencial uso de la información evaluativa (caso de UNO Habitat y PNUMA).
- ✓ La existencia de órganos de gobierno altamente comprometidos y concienciados con la importancia de la práctica evaluativa como herramienta para el aprendizaje puede materializarse en un gran empuje para conseguir un alineamiento entre el plan estratégico de la organización y los resultados de las evaluaciones y, por tanto, la generación de cambios a través de la evaluación (PNUD).

Tal y como se ha discutido anteriormente, ante un escenario donde se produce la combinación de alta capacidad de aprendizaje organizacional y baja institucionalización de la función de evaluación, cabe preguntarse: ¿La existencia de políticas y prácticas evaluativas es relevante? Los datos existentes nos hacen prever que, efectivamente, las políticas y prácticas evaluativas juegan un papel importante.

Así, si analizamos los resultados, observamos que PNUMA y UNO Habitat presentan una mayor solidez en las prácticas evaluativas que PNUD. Esto es fundamentalmente debido a dos causas. En primer lugar, PNUD realiza la primera revisión por pares en 2005, el mismo año del establecimiento de las normas y estándares UNEG. Por lo tanto, es natural que en aquel momento PNUD no hubiera desarrollado un marco normativo en materia evaluativa comparable al de aquellas agencias que se sometieron al proceso de revisión por pares algunos años más tarde. En segundo lugar, en general los programas con funciones de evaluación mayoritariamente descentralizada (PNUD o UNICEF) presentan mayores desafíos en la correcta aplicación de sus políticas evaluativas. Estos desafíos aparecen, en particular, vinculados con el proceso de elaboración del plan de evaluación y el desarrollo del sistema de seguimiento a las recomendaciones. En ambos procesos se requiere de un flujo de comunicación y coordinación entre las oficinas de campo/regionales y las oficinas centrales que dificulta la práctica evaluativa y que, en los casos de organizaciones con una función de evaluación centralizada, no se precisa.

De hecho, en el caso del PNUD podemos observar que son las evaluaciones corporativas en donde precisamente se produce una mayor aplicación efectiva con respecto a su planificación y seguimiento y, en consecuencia, las que han logrado un mayor impacto en cambios organizacionales y gozan de mayor credibilidad y apreciación. En las evaluaciones descentralizadas, la aplicación de las políticas de

planificación y seguimiento de las evaluaciones no ha sido tan efectiva hasta el momento (UNEG/UNDP 2013, pág. 7; Baasterl 2014, pág. 66).

De igual forma, la mejora en cuanto a relevancia y utilidad de las evaluaciones en PMA a raíz de la puesta en marcha de políticas y prácticas evaluativas sólidas (sistema de seguimiento a recomendaciones, sistema de control de calidad) como consecuencia de la aprobación de la Política de Evaluación en el 2008, sugieren un claro alineamiento entre ambos elementos (aplicación de políticas y prácticas evaluativas y uso de la evaluación) (UNEG/WFP 2014, pág. 48).



## 6. Conclusiones finales

En este último capítulo se presentan las conclusiones finales de la tesis. Para ello, tras presentar un breve resumen de lo descrito en las fases anteriores, se procede a describir las principales conclusiones del estudio empírico realizado a través del método QCA. Seguidamente, se revisan las hipótesis planteadas para proceder a la verificación y/o el rechazo de las mismas. Por último, la tesis finaliza con una identificación de los aspectos a tener en cuenta en futuros análisis de la capacidad evaluativa y, asimismo, se señalan algunos de los limitantes existentes en la actualidad para mejorar la capacidad evaluativa.

La presente tesis se ha centrado en el concepto de capacidad evaluativa. Hemos visto, a través de la revisión del marco teórico, el interés que este concepto ha generado y genera en la comunidad evaluadora, comunidad que, a lo largo de los años, ha realizado importantes aportes a la clarificación de este concepto y a la elaboración de modelos para determinar sus causas principales. Una gran parte del interés que ha suscitado la capacidad evaluativa está relacionada con la necesidad de comprobar hasta qué punto la práctica evaluativa añade valor a la organización y a sus resultados y, por tanto, amerita recursos para su fortalecimiento.

A pesar de los múltiples esfuerzos de investigación en torno a la capacidad evaluativa, la gran mayoría de autores y autoras inciden en la necesidad de seguir profundizando en esta área de estudio y, en particular, en la verificación empírica de los modelos conceptuales existentes (Suarez-Balcázar et al 2010). Esto es debido a que, hasta la fecha, la evidencia empírica existente es muy limitada y, en consecuencia, surgen cuestiones de gran interés sobre si las prácticas efectuadas para el desarrollo de la capacidad evaluativa en las organizaciones han incidido y contribuido realmente a mejorar el uso de la evaluación, a incrementar su calidad y, en definitiva, a introducir la práctica evaluativa como algo habitual dentro la actividad organizacional. Por lo tanto, esta tesis se ha llevado a cabo con el objetivo de aportar evidencia empírica al marco teórico existente, contribuyendo a su vez al desarrollo de nuevo conocimiento en la materia.

La tesis se centra en el marco del Sistema de NU, el cual ha mostrado a lo largo de los años un gran interés en la capacidad evaluativa como herramienta efectiva para el logro de resultados de desarrollo. Dicho interés ha sido expresado en la Declaración de París, el Programa de Acción de Accra y la Alianza de Busan para la Cooperación Eficaz al Desarrollo y ha culminado en la Resolución A/RES/69/237 del Segundo Comité de la Asamblea General de NU de Diciembre de 2014, a través de la cual, el año 2015 ha sido declarado Año Internacional de la Evaluación. Esta resolución resume los importantes esfuerzos puestos en marcha por el Sistema de NU para fortalecer la práctica evaluativa dentro de sus organismos, así como para promover el desarrollo de la capacidad evaluativa entre sus estados miembros como herramienta clave de la planificación estratégica y de la medición de resultados.

Sin embargo, tal y como concluyen recientes informes en la materia<sup>102</sup> provenientes de la Oficina de Servicios de Supervisión Interna, OIOS y de la Unidad de Inspección Conjunta, JIU (organismos del Sistema de NU que poseen el mandato de analizar, reportar y proponer medidas para la optimización de los recursos así como velar por el buen funcionamiento del Sistema), a pesar de los esfuerzos realizados en el marco del Sistema de NU, las evaluaciones no son utilizadas de manera plenamente efectiva en la toma de decisiones y su relevancia en el aprendizaje organizacional es aún limitada.

Por tanto, esta tesis trata de ahondar en las causas que explican la limitada utilización de la práctica evaluativa arriba mencionada. De esta manera, partiendo de la base de que existe un marco teórico que describe cuáles son los factores claves para el desarrollo de la capacidad evaluativa, la hipótesis de partida de esta tesis pone de manifiesto que es preciso diferenciar y priorizar unos factores explicativos frente a otros. Para ello, el estudio empírico aporta información y análisis sobre los factores más relevantes a la hora de apostar por el desarrollo de la capacidad evaluativa en aras a favorecer la calidad y el uso de las evaluaciones, tomando como referencia la función de evaluación en el marco del Sistema de NU.

Una de las principales limitaciones existentes en cualquier investigación empírica sobre el uso de la evaluación consiste en la dificultad de conseguir información verificable que permita analizar de manera concluyente si realmente se ha producido una utilización de la evaluación. En esta investigación, los informes de revisión por pares elaborados por paneles de evaluadores/as independientes y publicados por la UNEG para el análisis de las funciones de evaluación de los organismos, que voluntariamente acceden a someterse a dicho escrutinio, han sido la fuente de información principal. En concreto, se ha procedido al análisis de diez agencias del Sistema de NU, las cuales forman un grupo representativo de cómo se realiza la evaluación dependiendo de su ubicación dentro del organigrama del Sistema de NU.

El método utilizado es el Análisis Comparativo Cualitativo (QCA), el cual está orientado a aquellos casos que permiten realizar un análisis formal y sistemático de causalidad. Este método, desarrollado por Charles Ragin, fue concebido para facilitar el análisis empírico de muestras con un número de casos reducido. Surge principalmente con el objetivo de subsanar las importantes limitaciones del análisis cualitativo tradicional en esta materia, como la escasa posibilidad de generalización, la tendencia a enfocarse en exceso en una descripción particularista del caso de estudio concreto y la falta de transparencia y sistematicidad en el manejo de la información empírica (Ariza y Gandini 2012). A través del método QCA, es posible establecer los factores (o combinación de condiciones) suficientes para el logro de los resultados objeto de estudio.

Más específicamente, en esta tesis se han identificado tres condiciones que emanan del marco teórico como factores explicativos de la capacidad evaluativa: la institucionalización de la función de evaluación, la presencia de políticas y prácticas evaluativas, y la capacidad de aprendizaje de la entidad. Asimismo, se han estudiado en cada una de las diez agencias que componen la muestra, el efecto de estas

---

<sup>102</sup> JIU/REP/2014/6 y A/70/72

condiciones (de forma individual y combinada) sobre la consecución de los dos resultados esperados, resultados que también se derivan de la definición teórica comúnmente aceptada de capacidad evaluativa y que son: la calidad de las evaluaciones y el uso de la evaluación. El método QCA ha permitido en este caso identificar la combinación (o combinaciones) de condiciones más proclives a favorecer la calidad y el uso de las evaluaciones a partir de los factores mencionados, identificados en el marco teórico.

## **Conclusiones principales**

A continuación se presentan las principales conclusiones derivadas de la aplicación del método QCA para la identificación de relaciones de causalidad entre los factores determinantes de la capacidad evaluativa y sus resultados esperados.

1. Los factores identificados en el marco teórico son explicativos de la capacidad evaluativa.

Tal y como resumen Cousins et al (2014), los factores inherentes al desarrollo de capacidad evaluativa pueden englobarse dentro de dos categorías: la capacidad de hacer evaluaciones de calidad y la capacidad de hacer uso de las evaluaciones. El marco teórico descrito en esta tesis desgrana en profundidad los factores causales o condiciones dentro de cada una de estas dos categorías. Dichos factores han sido agrupados en el modelo propuesto en esta tesis en torno a tres condiciones: 1) Capacidad de aprendizaje, 2) Institucionalización de la función de evaluación y 3) Políticas y prácticas evaluativas.

El análisis realizado a través del método QCA confirma que las tres condiciones propuestas y que emanan del marco teórico son altamente explicativas de los resultados de capacidad evaluativa. Los altos índices de consistencia y cobertura resultantes en cada una de las condiciones, demuestran que los tres factores (capacidad de aprendizaje, institucionalización de la función de evaluación y políticas y prácticas evaluativas) son sólidos y determinantes para que se produzcan resultados. Lo anterior induce a pensar que definitivamente, el marco teórico recoge de manera efectiva los factores explicativos de la capacidad evaluativa.

2. Un marco normativo adecuado es fundamental para lograr resultados positivos en cuanto a la calidad de las evaluaciones se refiere; y en contraste, la capacidad de aprendizaje se convierte en un factor determinante para lograr el uso efectivo de la evaluaciones.

La definición propuesta por Stockdill et al 2002 y que sirve como base de referencia fundamental en el marco teórico existente sobre capacidad evaluativa, propone que la capacidad evaluativa persigue dos propósitos fundamentales: calidad de las evaluaciones y un uso rutinario de las mismas dentro de las organizaciones. Los resultados del análisis realizado a través de esta tesis

revelan que, a pesar de que ambos propósitos están altamente relacionados, el logro de la calidad y el logro de la utilización no tienen por qué ser obtenidos necesariamente a través de los mismos factores.

En este sentido, hay ciertos factores que favorecen de manera más efectiva resultados en calidad de la evaluación y otros factores que facilitan la utilización de la evaluación en la toma de decisiones y foros de referencia. Más específicamente, la existencia y aplicación de políticas y prácticas evaluativas efectivas se revela como una condición fundamental para el logro de una evaluación de calidad y, por ello, es una condición más importante que la institucionalización de la función de evaluación o la capacidad de aprendizaje organizacional a la hora de conseguir este resultado (Gibbs et al 2002; Preskill y Boyle 2008). Por otra parte, los resultados del análisis son concluyentes a la hora de determinar que la capacidad de aprendizaje es una condición fundamental para conseguir un uso rutinario de la práctica evaluativa en la toma de decisiones y en el aprendizaje organizacional. La institucionalización de la función de evaluación, a pesar de poseer una gran relevancia principalmente para favorecer la rendición de cuentas y la percepción de independencia e imparcialidad de la función de evaluación, no favorece necesariamente la utilización de la práctica evaluativa e incluso puede llevar a situaciones de aislamiento y distanciamiento de la función de evaluación con el resto de funciones dentro de la organización.

3. La relación entre calidad y uso debe ser tomada en cuenta para la puesta en marcha de estrategias coherentes de capacidad evaluativa.

Del análisis precedente de los informes de revisión por pares realizados para las agencias de la muestra se confirma que la calidad de la evaluación puede ser un factor influyente en el uso (UNEG/UNWomen, pág. 20). De hecho, se han podido observar algunas situaciones en las que los esfuerzos destinados a favorecer la calidad en las evaluaciones han derivado también en un mejor uso y una mayor apreciación de las mismas (Heider 2011).

Por citar un ejemplo concreto, la iniciativa llevada a cabo por la gerencia de UNICEF dirigida a incluir los resultados del sistema del control de calidad (GEROS), gestionados por la unidad de evaluación, como parte de los informes presentados a la Junta Ejecutiva ha tenido como consecuencia una mayor receptividad y apoyo hacia la evaluación por parte de la Junta Ejecutiva y del personal de UNICEF y, por tanto, una mayor relevancia de la evaluación en los foros de toma de decisiones (Universalía 2013, pág. v)<sup>103</sup>.

---

<sup>103</sup> Esto también puede producirse a la inversa, dado que una percepción negativa sobre la calidad de la evaluación pueda afectar su utilización. En el caso de la revisión por pares de ONUDI, algunas de las personas responsables de la gestión reportaron que una de las principales causas de la limitada utilización en el sistema de seguimiento a las recomendaciones de los informes de evaluación, era la escasa calidad de las recomendaciones (UNEG/UNIDO 2010, págs. 10, 26, 47-48).

Por lo tanto, a pesar de que la calidad de la evaluación no tenga necesariamente que implicar de forma directa una mayor utilización de la práctica evaluativa dentro de una organización, sí que ésta se constituye como un factor susceptible de mejorar la percepción existente de la evaluación como herramienta de mejora, lo que derivaría *a posteriori* en un mayor uso de la misma.

#### 4. Los factores que favorecen la calidad en las evaluaciones.

El estudio empírico realizado apunta a que la realización de evaluaciones de calidad se consigue principalmente a través del establecimiento de políticas y prácticas evaluativas que vayan acompañadas de sólidos esfuerzos orientados a fortalecer la capacidad de aprendizaje (a través de estrategias de disseminación de la información y de enfoques participativos) y/o por esfuerzos dirigidos a fortalecer la institucionalización de la función de evaluación (dotarle de una mayor independencia estructural y financiera dentro de los organismos). Los resultados del análisis permiten concluir que la calidad es un aspecto multidisciplinar, al depender de varias variables que van más allá de la puesta en marcha de políticas y prácticas sólidas que faciliten el rigor metodológico y la aplicación de procedimientos evaluativos. A pesar de que la presencia de políticas y prácticas evaluativas es fundamental para lograr resultados positivos en calidad (condición necesaria), dicho factor debe ir acompañado de otras medidas como, por ejemplo, la dotación adecuada de recursos financieros y personal, así como la necesidad de garantizar que la función de evaluación sea independiente de presiones internas o externas.

Combinar prácticas y políticas evaluativas sólidas con otras medidas adicionales es especialmente relevante en entidades donde la función de evaluación esta descentralizada. Las conclusiones de los estudios de revisión por pares coinciden en observar importantes debilidades en las funciones de evaluación descentralizadas que afectan a la calidad de las evaluaciones que producen.

Entre las principales carencias en un esquema descentralizado, se destacan los escasos recursos financieros y técnicos de que dispone la función de evaluación descentralizada, así como también el hecho de que las políticas de evaluación existentes en dichas organizaciones no clarifican suficientemente los roles, ni el nivel de apoyo que la unidad de evaluación central debe proporcionar a sus oficinas regionales y nacionales.

#### 5. Los factores que favorecen el uso de la evaluación.

La capacidad de aprendizaje se rebela como una condición fundamental y necesaria para el uso de la evaluación (Cousins et al 2004; GAO 2003; Owen 2003). En todos los casos estudiados, no es posible lograr que la evaluación sea utilizada si no es a través de una apuesta clara por la disseminación de la información y del conocimiento evaluativo, por la introducción de enfoques participativos y de la promoción de la cultura evaluativa a nivel organizacional.

Sin cultura de aprendizaje, el resto de factores difícilmente van a lograr por sí mismos la plena utilización de la función de evaluación.

Así, deficiencias en la cultura de aprendizaje producen efectos negativos en el logro de los resultados que persigue la capacidad evaluativa, a pesar de que las otras condiciones del marco teórico estén presentes (estructuras, recursos y políticas/prácticas evaluativas). La cultura de aprendizaje es, por tanto, un eje central en la capacidad evaluativa y su presencia permite que los demás factores facilitadores sean efectivos.

El estudio permite concluir que el uso de la evaluación es más efectivo en las organizaciones cuando además de la apuesta por el aprendizaje (condición necesaria), la organización dedica esfuerzos al establecimiento de un sólido marco normativo de políticas y prácticas evaluativas, en el que se delimiten con claridad los roles y las responsabilidades inherentes al ejercicio de la evaluación, así como los procedimientos necesarios para su planificación y posterior seguimiento de sus recomendaciones.

6. El aprendizaje y el marco normativo como combinación efectiva para lograr resultados positivos en la capacidad evaluativa.

La estrategia de combinar la capacidad de aprendizaje y el establecimiento de un marco normativo (políticas y prácticas evaluativas) se presenta en el estudio empírico como la combinación más efectiva de factores para conseguir simultáneamente resultados positivos tanto en calidad como en utilización de la evaluación. Así, cuando la capacidad de aprendizaje (entendida ésta como la habilidad de las organizaciones de aplicar el conocimiento adquirido en su práctica rutinaria) se complementa con el desarrollo y aplicación de un exhaustivo marco normativo de políticas y prácticas evaluativas, la combinación de ambas condiciones permite que los propósitos inherentes a la capacidad evaluativa (calidad y uso) se realicen de manera más consistente y efectiva.

La combinación de las condiciones de aprendizaje y marco normativo se observa como la combinación más frecuente sobre todo en aquellas organizaciones que buscan el equilibrio entre el aprendizaje y la rendición de cuentas.

El desbalance bien sea hacia al aprendizaje o hacia la rendición de cuentas puede llevar a una aplicación limitada y no efectiva de los esfuerzos destinados al desarrollo de capacidad evaluativa. Así, cuando el empeño por fortalecer la evaluación en las organizaciones se orienta fundamentalmente a uno de los dos propósitos principales de la evaluación, bien sea la rendición de cuentas o el aprendizaje, se observa que los logros conseguidos en cuanto al desarrollo de

capacidad evaluativa se refiere se ven mermados y tienden a favorecer únicamente a uno de los dos resultados posibles (calidad o uso)<sup>104</sup>.

#### 7. La posible sobredimensión de la institucionalización de la función de evaluación.

El estudio empírico realizado permite entrever que, entre los tres factores considerados (institucionalización, capacidad de aprendizaje y políticas/prácticas evaluativas), el factor de institucionalización de la función de evaluación posiblemente sea el menos determinante para lograr resultados en calidad y uso. Así, la institucionalización no forma parte de las condiciones necesarias para ninguno de los dos resultados y únicamente se encuentra en una de las dos soluciones derivadas de la tabla de la verdad para lograr resultados satisfactorios en calidad, siempre que el marco normativo (políticas y prácticas evaluativas) esté presente.

Además, también se observa que la ausencia o un bajo nivel de institucionalización de la función de evaluación ha llevado incluso a favorecer la utilización de las evaluaciones en algunos casos. Así, se han encontrado casos de organizaciones que, a pesar de contar con una débil institucionalización de la función de evaluación (limitada independencia estructural, o pocos recursos financieros destinados al ejercicio de la evaluación), su apuesta por una cultura de aprendizaje ha favorecido la presencia y el uso de la evaluación en la toma de decisiones. Los resultados dejan entrever que una mayor cercanía de la función de evaluación con las unidades programáticas ha podido ser un elemento fundamental a la hora de repercutir positivamente en una apreciación por parte del personal de la organización hacia el mandato de la unidad de evaluación, incrementando así su compromiso con el aprendizaje.

Por el contrario, las organizaciones que cuentan con una fuerte institucionalización de la función de evaluación corren el riesgo de aislarse del resto de la organización, especialmente cuando priorizan sus esfuerzos a favor de la rendición de cuentas frente al aprendizaje. La institucionalización de la función de evaluación, cuando se hace principalmente a través de la creación de estructuras de unidades de evaluación independientes pero sin fortalecer el componente del aprendizaje organizacional, no facilita el uso y la sostenibilidad de la práctica evaluativa. Al contrario, tiende a aislarla e incomunicarla del resto de las funciones, fomentando un carácter de vigilancia y control en vez de cooperación y aprendizaje.

---

<sup>104</sup> Así, la falta de rendición de cuentas puede también ser determinante en un mayor abandono en los controles de calidad, principalmente en la evaluación descentralizada (UNICEF, Baasterl 2014, pág. 67). Por su parte, cuando la orientación hacia la rendición de cuentas es primordial y el aprendizaje cobra un papel muy secundario, se pueden llegar a situaciones de cierto aislamiento que dificultan que los resultados de la evaluación sean compartidos y utilizados (UNEG/GEF 2009, pág. 11).

## 8. El papel de los órganos de gobierno en la evaluación.

Las agencias objeto de estudio han mostrado una gran diversidad con respecto al rol y responsabilidades de sus respectivos órganos de gobierno en su política de evaluación. Los resultados del estudio muestran que la participación de los órganos de gobierno, principalmente en tareas de aprobación y reporte de políticas e información evaluativa, es importante para incrementar la apreciación y el respeto del personal de la organización en general hacia el trabajo de las unidades de evaluación.

La falta de relevancia de la función de evaluación muchas veces va acompañada de escasas oportunidades por parte de la persona responsable de la unidad de evaluación de comunicar directamente los resultados de las evaluaciones a los órganos de gobierno. La cercanía entre la unidad de evaluación y los órganos de gobierno ayuda a generar una percepción positiva de la evaluación en la organización como agente de cambio y referencia clave en la toma de decisiones (Gibbs et al 2002, Mackay 2002).

Sin embargo, el establecer canales de comunicación fluida entre las unidades de evaluación y los órganos de gobierno no implica necesariamente que estos últimos se conviertan en el usuario intencionado de la evaluación. Una atención desbalanceada de la unidad de evaluación en favor de los órganos de gobierno, y dejando en un segundo plano a las unidades programáticas como usuarios intencionados, puede llevar a tensiones y confusión en la organización a nivel interno que, de nuevo, derive en una priorización de los propósitos de rendición de cuentas con respecto al aprendizaje.

### **Verificación de la hipótesis**

La hipótesis a contrastar en esta tesis es que si el fomento de la cultura de aprendizaje viene acompañado del establecimiento y de la aplicación de un sólido marco normativo, entonces la combinación de ambas condiciones se convierte en un vehículo efectivo para el logro de la capacidad evaluativa necesaria en las instituciones de las NU, que permita no solo la mejora de la calidad de las evaluaciones sino también su influencia en la toma de decisiones. Esta hipótesis ha sido contrastada y verificada en el análisis empírico realizado a través de QCA.

A pesar de que el marco teórico incluye a la cultura de aprendizaje como uno de los factores que favorecen la capacidad evaluativa, los efectos de su combinación con la existencia de un marco normativo consistente no están reconocidos en los modelos teóricos que actualmente analizan la capacidad evaluativa. La investigación realizada a través de este trabajo ha permitido comprobar que la cultura de aprendizaje se constituye en una condición necesaria para lograr resultados satisfactorios en el ámbito de la utilización de la evaluación. Si la cultura de aprendizaje, es combinada con un marco normativo adecuado, se convierte entonces en una fórmula efectiva para

producir, a su vez, no solo buenos resultados en cuanto al uso sino también en relación a la calidad de las evaluaciones en las organizaciones, cumpliendo así por tanto los objetivos inherentes a los esfuerzos de capacidad evaluativa (calidad y utilización).

Sin cultura de aprendizaje organizacional, el marco normativo y la institucionalidad de la función de evaluación son factores que pierden efectividad y difícilmente van a lograr por sí mismos la plena utilización de la función de evaluación. Así, deficiencias en la cultura de aprendizaje organizacional producen efectos negativos en el logro de los resultados que persigue la capacidad evaluativa, a pesar de que las otras condiciones del marco teórico estén presentes (especialmente las relativas a fortalecer la institucionalización de la función de evaluación). La cultura de aprendizaje organizacional es, por tanto, un eje central en el desarrollo pleno de la capacidad evaluativa y su presencia facilita que los demás factores sean efectivos.

En consecuencia, la práctica evaluativa, aunque presente en las instituciones del Sistema de NU, no está todavía asentada como herramienta de aprendizaje organizacional, siendo la capacidad evaluativa de las organizaciones aún débil, hecho que limita la influencia de las evaluaciones en la toma de decisiones.

En cualquier caso, la evaluación en NU ha experimentado grandes avances en cuanto a su consolidación como disciplina y en cuanto a la dotación de recursos para su establecimiento organizativo y programático. De esta forma, se han fortalecido las unidades de evaluación como estructuras funcionalmente independientes con el fin de proporcionar evaluaciones imparciales y de calidad. Otros aspectos como el de la integración de la evaluación como parte del aprendizaje organizacional están aún pendientes. Debido a ello, las evaluaciones siguen teniendo un carácter predominantemente aislado y este hecho limita su utilidad y su influencia.

Entre los diferentes enfoques aplicados por las organizaciones de UN, se pueden distinguir dos principalmente: las organizaciones que han priorizado el fortalecimiento de la función de evaluación desde su institucionalidad con la rendición de cuentas como propósito principal, frente a las organizaciones que han primado los enfoques consultivos, participativos y donde el aprendizaje tiene un valor fundamental en la evaluación. Este segundo grupo ofrece mejores resultados principalmente en cuanto a la sostenibilidad de la función de evaluación, ya que la institucionalidad de la función de evaluación en sí misma, no lleva necesariamente a fortalecer una mayor utilización de las evaluaciones.

Los organismos con una mayor capacidad de aprendizaje organizacional y que han combinado este factor con el establecimiento de un marco normativo sólido, a través de exhaustivas políticas y prácticas evaluativas, muestran una mayor capacidad evaluativa, tanto en cuanto a la calidad como en el uso sostenible de las evaluaciones. Existe una mayor evidencia de que, en este contexto, la evaluación de calidad forma parte de la toma de decisiones en materia programática (programas evaluados) así como en relación a las políticas institucionales.

## Conclusión final

Las conclusiones de la tesis, anteriormente expuestas, permiten verificar la hipótesis formulada al comienzo de esta investigación, así como dar respuesta a las interrogantes planteadas. La capacidad evaluativa se constituye como el fin de un proceso en el que las premisas teóricas en cuanto a lo que constituye capacidad evaluativa y en relación a cuáles son sus causas explicativas han sido ratificadas mediante el análisis empírico.

Además, los resultados de la investigación sugieren que el hecho de establecer una cierta priorización dentro de los factores determinantes de la capacidad evaluativa puede resultar en un ejercicio eficaz y útil para los organismos. En particular, promover el aprendizaje organizacional junto con el desarrollo de un marco normativo basado en políticas y prácticas evaluativas facilita una mayor optimización de los resultados esperados sobre capacidad evaluativa en cuanto a calidad y uso de las evaluaciones. A su vez, los resultados del estudio empírico nos ayudan a entender las causas por las que los esfuerzos destinados hasta la fecha a desarrollar capacidad evaluativa en muchas organizaciones no han permitido una utilización plena de la evaluación en la toma de decisiones.

Una de las principales lecciones aprendidas y obtenidas a través de este estudio se centra en la idea de que el fortalecimiento institucional de la función de evaluación no necesariamente conlleva a un mayor uso de la evaluación en las organizaciones. La importancia otorgada a conseguir unidades de evaluación independientes estructuralmente, y con capacidad para planificar y ejecutar sus actividades, debe ir acompañada de una apuesta clara por hacer de las NU un sistema de organizaciones con capacidad de aprender a partir del tipo de información que la evaluación puede proporcionar. La capacidad de aprendizaje de las organizaciones es lo que realmente puede conseguir que la evaluación se convierta en elemento generador de cambio.

Sin embargo, también es verdad que el contexto actual no favorece claramente la elección de esta estrategia. Ante un contexto internacional de recursos limitados y donde cada vez más la cooperación al desarrollo va intrínsecamente ligada a una mayor demanda y presión dirigida a la obtención de resultados, es fácil entender que en relación al fortalecimiento de las funciones de evaluación se hayan primado aquellos aspectos más favorables a garantizar la rendición de cuentas frente al aprendizaje. Este desequilibrio entre los dos propósitos naturales de la evaluación dificulta que la práctica evaluativa evolucione en la dirección adecuada, que le permita convertirse en una herramienta clave para la generación de conocimiento organizacional.

La práctica evaluativa como herramienta de conocimiento requiere que las evaluaciones sean concebidas como partes integrantes de un sistema de información donde sus datos, hallazgos y conclusiones fluyan con libertad dentro de la organización y se complementen con información sustantiva proveniente de otras fuentes, como por ejemplo la que emana de los sistemas de monitoreo de los proyectos y programas. Así,

la revisión del marco teórico ha puesto de relevancia que, a pesar de sus comienzos dificultosos, la evaluación y el monitoreo no solo pueden convivir de forma armónica como disciplinas afines dentro de las organizaciones, sino que además pueden retroalimentarse y complementarse en aras a favorecer la calidad de la información que generan y consecuentemente el flujo de la misma, con los efectos positivos que lo anterior conlleva en relación al aprendizaje organizacional.

Sin embargo, en todos los informes de revisión por pares analizados en esta investigación, se ha podido constatar que la función de monitoreo tiene un carácter débil en el Sistema de NU y, ante la falta de calidad y sincronización del monitoreo con la evaluación, la evaluación en muchos casos tiende a centrarse en aspectos más propios del monitoreo operativo que de relevancia estratégica. Otorgar a la evaluación un carácter estratégico tiene una importancia crucial en relación al logro del resultado relacionado con el uso de las evaluaciones. En el contexto actual, ante la cada vez más frecuente tendencia a centrar las evaluaciones en aspectos técnicos y operativos frente a aspectos estratégicos y de largo alcance, difícilmente se puede conseguir que la práctica evaluativa se posicione como agente clave de mejora en la toma de decisiones estratégicas. En línea con Stufflebeam y Shinkfield (1985), "*el principal propósito de la evaluación no es probar sino mejorar*" (pág. 151). Así a la tradicional pregunta de ¿estamos haciendo lo adecuado?, se suman las preguntas de ¿lo estamos haciendo bien? y ¿qué se puede hacer mejor?. En este sentido, evaluación y monitoreo deben actuar como flujos complementarios ya que las deficiencias en la función de monitoreo difícilmente pueden ser subsanadas de forma satisfactoria para la organización a través de la evaluación.

Por lo tanto, podemos afirmar que todavía queda camino por recorrer para que la evaluación sea en definitiva una práctica rutinaria y efectiva en las organizaciones del Sistema de NU y forme parte del flujo de información que nutra el aprendizaje dentro de las mismas. Sin embargo, este hecho también pone de manifiesto la necesidad de seguir investigando y aportando propuestas con respecto a cómo mejorar la capacidad evaluativa, con el fin de que la práctica evaluativa se convierta en una herramienta de cambio que permita la optimización de los recursos disponibles y, en última instancia, el logro de los objetivos de desarrollo inherentes al Sistema de NU.



## Bibliografía

- AGCI (2011). Cuarto Foro de alto nivel sobre la eficacia de la ayuda. *Centro de Documentación*.
- Alaimo, S.P. (2008). Nonprofits and evaluation: Managing expectations from the leaders's perspectiva. En *J.G. Carman y K.A Fredericks Eds. Nonprofits and Evaluation. New Directions for Evaluation*, 119, 73-92.
- Alcocer, J.; Lizárraga, P., Delgadillo, J. (1997). *Sondeo Sobre Utilización de Métodos Participativos de Monitoreo y Evaluación en Bolivia*. Presentado en el Taller Internacional sobre Monitoreo y Evaluación participativa: Experiencias y lecciones. Cavite, Philippines, 24–29 Nov 1997.
- Alkin, M.C y Coyle, K. (1988). Thoughts on evaluation utilization, misutilization and non-utilization. *Studies in Educational Evaluation*, 14, 331-340.
- Alkin, M.C. (1975). Evaluation: Who needs it? Who cares? *Studies in Educational Evaluation*, 1, 201-212.
- Alkin, M.C. (1990). *Debates on evaluation*. Newbury Park, CA, Sage.
- Alkin, M.C. (2004). *Evaluation roots: tracing theorists' views and influences*. Thousand Oaks, CA: Sage.
- Alkin, M.C. (2005). Utilization of evaluation. En *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage.
- Alonso, J.A., y Mosley, P. (1999). *La eficacia de la cooperación internacional al desarrollo: Evaluación de la ayuda*. Biblioteca Civitas, Economía y Empresa. Colección Economía.
- Amo, C., y Cousins, J. B. (2007). Going through the process: An examination of the operationalization of process use in empirical research on evaluation. *New Directions for Evaluation*, 116, 5-26.
- Argyris, C. y Schön, D. (1978). *Organisational learning: A theory of action perspective*. Reading, Mass: Addison Wesley.
- Ariza, M. Gandini, L. (2012). El análisis comparativo cualitativo como estrategia metodológica. En Ariza, M., Velasco, L. (Eds.). *Métodos cualitativos y su aplicación empírica: Por los caminos de la investigación sobre migración internacional* (México D.F: Universidad Nacional Autónoma de México. Instituto de Investigaciones Sociales Tijuana B.C. El Colegio de Frontera Norte.
- Arnold, M. (2006). Developing evaluation capacity in extension 4-H field faculty. *American Journal of Evaluation*, 27 (2), 257-269.
- Ashworth, K., Greenberg, D. and Walker, R. (2004). Meta-evaluation: discovering what works best in welfare provision, *Evaluation*, 10,1.
- Ayers, T.D. (1987). Stakeholders as partners in evaluation: A stakeholder-collaborative approach. *Evaluation and programme planning*, 10, 263-271.

- Baastel (2014). Review of the UNDP Evaluation Policy. <http://web.undp.org/evaluation/>
- Baizerman, M., Compton, D. W., y Stockdill, S.H. (2002). New directions for ECB. *New Directions for Evaluation*, 93, 109-119.
- Ballart, X. (2003). Evaluación de políticas (Marco conceptual y organización institucional). *Revista de Asuntos Políticos* abril-junio.
- Bamberger, M. (2000). The evaluation of international development programs: A view from the front. *American Journal of Evaluation*, 21 (1), 95- 102.
- Bastoe, O. (2006). Implementing Results Based Management. En Rist, R., y Stame, N. Eds. *From studies to streams: Managing evaluative systems*. Comparative policy evaluation.
- Binnendijk, A. (2000). Results Based Management in the Development Cooperation Agencies: A review of experience. DAC Working Party on Aid Evaluation.
- Birkeland, S., Murphy-Graham, E., Weiss, C. (2005). Good reasons for ignoring good evaluation: The case of the drug abuse resistance education program. *Evaluation and program planning* 28, 247-256.
- Blalock, A.B. (1999). Evaluation Research and the Performance Management Movement. *Evaluation*, 5 (2),117-149.
- Bourgeois, I., y Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation*, 34 (3), 299–319.
- Boyle, R. (1999). Professionalizing the evaluation function: Human resource development and the building of evaluation capacity. En R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity* (135–151). New Brunswick, NJ: Transaction
- Boyle, R., Lemaire, D., y Rist, R. C. (1999). Introduction: Building evaluation capacity. En R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity* (1-19). New Brunswick, NJ: Transaction Publishing.
- Brooks-Harris, J. E., y Stock-Ward, S. R. (1999). *Workshops: Designing and facilitating experiential learning*. Thousand Oaks, CA: Sage.
- Bruner, J. S. (1971). *The Relevance of Education*. London: Penguin Education.
- Bustelo, M. (2004). ¿Cómo pueden contribuir el proceso de creación de códigos normativos al desarrollo de una cultura de evaluación? Elementos para el debate en el caso español. *GAPP*, No 30-31. Mayo-Diciembre 2004.
- Campbell D. (1998). *Writing Security: United States foreign policy and the politics of identity*. Manchester: Manchester University Press.
- Caracelli, V.J. (2000). Evaluation use at the threshold of the twenty-first century. En V. J. Caracelli y H. Preskill (Eds.), the expanding scope of evaluation use. *New Directions for Evaluation*, 88, 99-111. San Francisco. CA: Jossey-Bass.
- Carlsson, J.M, Fallenius A. y Lovgren, E. (1999). Are evaluations useful? Cases from Swedish Development Cooperation. Stockholm: Sida.

- Carman, J.C. (2007). Evaluation practice among CBOs. *American Journal of Evaluation*, 28, 60-75.
- Carman, J.C., y Fredericks, K.A. (2008). Nonprofits and evaluation. *New Directions for Evaluation*, 119, San Francisco. CA: Jossey-Bass
- Carman, J.C., y Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half empty or half-full? *American Journal of Evaluation*, 31(1), 84-104.
- Cioffi-Revilla, C. (1981). Fuzzy sets and models of international relations. *American Journal of Political Science* 25 (1), 129-59.
- CIPPEC (2012). Monitoreo y evaluación de políticas, programas y proyectos sociales. CIPPEC, UNICEF Argentina.
- Clarke, A. (1999). *Evaluation research: An introduction to principles, methods and practice*. London: Sage.
- Cohen, E. y Franco, R. (2005). Gestión Social. Cómo lograr eficiencia e impacto en las políticas sociales. México: CEPAL y Siglo Veintiuno.
- Collier, D (1998). Comparative method in the 1990s. Newsletter of the Comparative Politics Organized Section. *American Political Science Association* 9, 1-5.
- Compton, D.W., Glover-Kudon, R., Smith, I.E., y Avery, E. (2002). Ongoing capacity building in the American Cancer Society (ACS) 1995-2001. En D.W. Compton, M. Baizerman, y S.H. Stockdill (Eds.), *The Art, Craft and Science of Evaluation Capacity Building*. *New Directions for Evaluation*, 93 (47-61). San Francisco, CA: Jossey-Bass.
- Comstock, D. E., and Fox, R. (1993). Participatory Research as Critical Theory: The North Bonneville, USA Experience. En P. Park, M. Brydon-Miller, B. Hall, and T. Jackson (eds.), *Voices of Change. Participatory Research in the United States and Canada*. Toronto: OISE.
- Connolly, P., y York, P. (2002). Evaluating capacity-building efforts for nonprofit organizations. *OD Practitioner*, 34 (4), 33-39.
- Costello, B., Muller, L., Amitani, S. y Edmonds, E.A. (2005). Understanding the Experience of Interactive Art: Iamascope in Beta Space. En *Interactive Entertainment, Creativity & Cognition Studios Press*, 49-56.
- Cousins J. B. (1996). Consequences of researcher involvement in participatory evaluation. *Studies in Educational evaluation*, 22 (1), 3-27.
- Cousins, B. y K. Leithwood. (1986). Current Empirical Research on Evaluation Utilization. *Review of Educational Research* 56 (3), 331-364.
- Cousins, J. B y Earl L.M. (1992). The case for participatory evaluation. *Educational evaluation and policy analysis*, 14 (4), 397-418.
- Cousins, J. B y Earl L.M. (1995). Participatory evaluation in education. *Studies in evaluation use and organizational learning*. London. Falmer Press.

- Cousins, J. B., y Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 80, 5-23.
- Cousins, J. B., Goh, S., Clark, S., y Lee, L. (2004). *Integrating consequences of evaluation into evaluation capacity building enquiry*. Presentado en la reunión anual de la Asociación Americana de Evaluación, Baltimore, MD.
- Cousins J. B y Shulha L.M. (2006). A comparative analysis of evaluation utilization and its cognate fields of inquiry: current issues and trends. En *the Sage Handbook of Evaluation: Policies, programs and practices*, (266-291). Thousand Oaks, CA: Sage.
- Cousins, J. B., Goh, S. C., Elliott, C., y Aubry, T. (2008). *Government and voluntary sector differences in organizational capacity to do and use evaluation*. Presentado en la reunión anual de la Asociación Canadiense de Evaluación, Quebec.
- Cousins, J. B., Chouinard, J. A. (2012). *Participatory Evaluation Up Close: A Review and Integration of Research-Based Knowledge*. Greenwich, Connecticut: Information Age Publishing Inc.
- Cousins, J. B., Whitmore, E., y Shulha, L. (2013). Arguments for a common set of principles for collaborative inquiry in evaluation. *American Journal of Evaluation*, 34, 7–22.
- Cousins, J.B., Goh, S.C., Elliott, C.J. y Bourgeois, I. (2014). Framing the capacity to do and use evaluations. *New Directions for Evaluation*, 141, 7-23.
- Cronbach, L. J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., Walker, D.F., y Weiner, S.S. (1980). *Towards reform of program evaluation*. San Francisco, CA: Jossey Bass.
- Cummings, R. (1999). *Stakeholder involvement: A foundation for evaluation use*. Presentado en la reunión anual de la Asociación Americana de Evaluación. Orlando, FLA.
- Cummings, R. (2002). *Rethinking Evaluation Use*. Presentado en la reunión anual de la Asociación Austral Asiática de Evaluación. Australia.
- Charlton, B. G. (2002). Audit, accountability, quality and all that: The growth of managerial technologies in UK universities. En S. Prickett & P. Erskine-Hill (Eds.), *Education! education! education!- Managerial ethic sand the law of unintended consequences* (13–28). Thorverton, UK: Imprint Academic.
- Chaskin, R. J. (2001). Community capacity: A definitional framework and implications from a comprehensive community initiative. *Urban Affairs Review*, 36 (3), 291-323.
- Chelimsky E. (1997). The coming transformation. En E. Chelimsky y W. Shadish Eds. *Evaluation for the 21<sup>st</sup> century*. Thousand Oaks, CA. Sage.
- Chen, H. (1996). A comprehensive typology for program evaluation. *Evaluation Practice*, 17 (2), 121–130.

- Christie, C. (2007). Reported influence of evaluation data on decision makers' actions: An empirical examination. *American Journal of Evaluation* 28 (1), 8-25
- DAC/UNEG (2007). *Framework for Professional Peer Reviews*. The DAC/UNEG Joint Task Force on Professional Peer Reviews of Evaluation Functions in Multilateral Organizations. <http://www.uneval.org/document/detail/945>
- Daigneault, P.-M., y Jacob, S. (2009) Toward Accurate Measurement of Participation: Rethinking the Conceptualization and Operationalization of Participatory Evaluation, *American Journal of Evaluation*, 30 (3) 330-348.
- Davies, I. C., (1999). Evaluation and Performance Management in Government. *Evaluation*, 5 (2), 150-159. Sage.
- Davis, B. G. (2001). *Tools for teaching*. San Francisco: Jossey-Bass.
- De Lancer Julnes, P. (2006). Performance Measurement. *Evaluation* 12 (2), 219-235.
- Declaración de París <http://www.oecd.org/dataoecd/53/56/34580968.pdf>
- Duffy, B.P. (1994). Use and abuse of internal evaluation. En C.J. Stevens y M. Dial eds, *New directions for program evaluation: Preventing the misuse of evaluation*, 64 (25-32). San Francisco: Jossey-Bass.
- Duffy, J.L., y Wandersman, A. (2007). A review of research on evaluation capacity-building strategies. Presentado en la reunión anual de la Asociación Americana de Evaluación. Baltimore, MD.
- Edwards, R.W., Jumper-Thurman, P., Plested, B.A., Oetting, E.R y Swanson, L. (2000). *Community readiness: Research to practice*. *Journal of Community Psychology*, 28 (3), 291-307.
- Eggen, Paul D. y Kauchak, D.P. (1999). *Estrategias docentes. Enseñanza de contenidos curriculares y desarrollo de habilidades de pensamiento*. Fondo de Cultura Económica de Argentina. Buenos Aires.
- Executive Edge. (1998). *Strategic planner*. Executive Edge magazine, 16–17.
- Fals-Borda, O. (1980). Science and the Common People. Presentado en el Foro Internacional sobre Investigación Participativa, Ljubljana, Yugoslavia.
- Fals-Borda, O., and Anisur-Kahman, M. (1991) *Action and Knowledge. Breaking the Monopoly with Participatory Action Research*. New York: Apex Press.
- Fawcett, S.B., White, G.W., Balcazar, F.E., Suarez-Balcazar, Y., Mathews, R.M., Paine-Andrews, A., et al (1994). A contextual-behavioral model of empowerment: Case studies involving people with physical disabilities. *American Journal of Community Psychology*, 22, 471-496.
- Fawcett, S. B., Paine-Andrews, A., Francisco V.T., Schultz, J.A, Ritcher, K.P., Lewis, R.K., et at (1996). Empowering community health initiatives through evaluation. En D. Fetterman, S. Kaftarian y A. Wandersman Eds., *Empowerment Evaluation: Knowledge and Tools for self assessment and accountability* (256-276). Thousand Oaks, CA: Sage.

- Fawcett, S. B., Boothroyd, R., Schultz, J.A., Francisco, V.T., Carson, V., y Bremby R. (2003). Building capacity building for participatory evaluation within community initiatives. En Y. Suarez-Balcazar y G. Harper Eds, *Empowerment and participatory evaluation of community interventions*, 21-36, Binghamton, NY, Haworth Press.
- Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation practice*, 15 (1), 1-15
- Fetterman, D. M. (2001). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.
- Fetterman, D. M. (2005). En D. M. Fetterman and A. Wandersman, Eds. *Empowerment evaluation principles in practice* (1-26). New York, NY: Guilford.
- Fetterman, D. M. (2007). Empowerment evaluation: An exercise in efficiency. *From Policies to Results: Developing Capacities for Country Monitoring and Evaluation Systems*. UNICEF.
- Fetterman D, M, Kaftarian, A.J., y Wandersman, A. Eds. (1996). *Empowerment evaluation. Knowledge and tools for self assessment and accountability*. Thousand Oaks, CA. Sage publications.
- Fetterman, D. M., y Wandersman, A. Eds. (2005). *Empowerment evaluation principles in practice*. New York, NY: Guilford.
- Fetterman, D., y Wandersman, A. (2007). Empowerment evaluation: Yesterday, today, and tomorrow. *American Journal of Evaluation*, 28, 179–198.
- Fetterman, D., Rodríguez-Campos, L., Wandersman, A., y O’Sullivan, R. (2014). Collaborative, participatory and empowerment evaluation: Building a strong conceptual foundation for stakeholder involvement approaches to evaluation. *American Journal of Evaluation*, 35, 144–148.
- Feuerstein, M. T. (1986). *Partners in evaluation: Evaluating development and community programs with participants*. London, England: Macmillan.
- Fitzpatrick, J. L. (1998). Dialogue with Marsha Mueller. *American Journal of Evaluation*, 19, 87-99.
- Flaspohler, P., Duffy, J., Wandersman, A., Stillman, L., y Maras, M (2008). Unpacking prevention capacity: An intersection of research-to-practice models and community-centered models. *American Journal of Community Psychology*, 41 (3-4), 182-186.
- Forss, K; Rebien, CC y Carlsson, J. (2002). Process use of evaluations: Types of use that precede lessons learned and feedback. *Evaluation*, 8 (1), 29-45.
- Freire, P. (1970). *Pedagogy of the Oppressed*. New York: Seabury Press.
- Frey, K., y Ledermann, S. (2010). Introduction: Evidence-based policy: A concept in geographical and substantive expansion. *German Policy Studies*, 6 (2), 1-15,187.
- Fruchter, N., Cahill, M. and Wahl, E. (1998). *Building to Local Capacity: A Review of Technical Assistance Strategies*. New York University Institute for Education and Social Policy.

- Furubo, J. (2006). Why evaluations sometimes can't be used and why they shouldn't. En *From Studies to Streams*, R Rist and N. Stame. Transaction Publishers.
- Furubo, J.E. y Sandahl, R. (2002). Introduction – A Diffusion Perspective on Global Developments in Evaluation. En J. E. Furubo, R. Rist, and R. Sandahl (eds.), *International Atlas of Evaluation*, New Brunswick, NJ: Transaction Publishers.
- García-Iriarte, E., Suarez-Balcazar, Y., Taylor-Ritzler, T., y Luna, M. (2011). A catalyst for change approach to evaluation capacity building. *American Journal for Evaluation*, 32 (2), 162-182.
- Garvin, D.A. (1993). Building a learning organization. *Harvard Business Review*, July–august, 78–91.
- Gibbs, D., Napp, D., Jolly, D., Westover, B., y Uhl, G. (2002). Increasing evaluation capacity within community based HIV prevention programs. *Evaluation and Program Planning*, 25, 261-269.
- Goering, P. N y Wasylenki D.A (1993). Promoting the utilization of outcome study results by assuming multiple roles within an organization. *Evaluation and program planning*, 16, 329-334.
- Goh, S. C and Richards, G. (1997). Benchmarking the learning capability of organizations, *European Management Journal*, 15 (5), 575- 583.
- González, B., López, C., y Vidal, M. M. (2012). Estudio de las decisiones de internacionalización empresarial desde una perspectiva cualitativa: Aplicación del análisis comparativo cualitativo de categorías difusas. *Instituto valenciano de investigaciones económicas. Series EC. WP-EC-2012-08*.
- Goodman, R.M., Speers, M. A., Mcleroy, K., Fawcett, S., Kegler, M., Parker, E. et al (1998). Identifying and defining the dimensions of community capacity to provide a basis for measurement. *Health Education and Behavior*, 25 (3), 258-278.
- Government Accountability Office. (2003). *Program evaluation: An evaluation culture and collaborative partnerships help build agency capacity* (GAO-03-454). Washington, DC: Author.
- Greene J. (1988a). Stakeholder participation and utilization in program evaluation. *Evaluation review*, 12 (2), 91-116.
- Greene J. (1988b). Communication of results and utilization in participatory program evaluation. *Evaluation and programme planning*, 11, 341-355.
- Greene, J. C. (1990). Technical quality versus user responsiveness in evaluation practice. *Evaluation and program planning*, 13, 267-274.
- Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice*, 18, 25–35.
- Greene J. (1999). The Inequality of Performance Measurements. *Evaluation*, 5 (2), 160-172.
- Greene, J. C. (2000) Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation. *New Directions for Evaluation*, 85, 13–26.

- Greene, J., Lincoln, Y., Mathison, S., Mertens, D., y Ryan, K. (1998). Advantages and challenges of using inclusive evaluation approaches in evaluation practice. *American Journal of Evaluation*, 19 (1), 101–122.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., y Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly*, 82 (4), 581-629.
- Guarino, C., Santibañez, L., Daley, G., y Brewer, D. (2004). *A review of the research literature on teacher recruitment and retention*. Technical report TR-164-EDU. Santa Monica, CA.
- Guba, E.G y Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park, CA. Sage.
- Heider, C. (2007). Evaluating policy. A perspective from multilateral experience. En *From Policies to Results: Developing Capacities for Country Monitoring and Evaluation Systems*. UNICEF.
- Heider, C. (2011). Conceptual framework for developing evaluation capacities: building on good practice. En Rist, R., Boily, M.-H., y Martin, F. Eds. *Influencing change: Evaluation and Capacity Building*. The World Bank.
- Henry, G. T. (2002). Why not use? En V J Caracelli and H Preskill (eds) *The Expanding Scope of Evaluation Use*, New Directions for Evaluation 88, San Francisco: Jossey-Bass.
- Henry, G. T., y Mark, M. M. (2003a). Toward an agenda for research on evaluation. *New Directions for Evaluation*, 97, 69-80.
- Henry, G. T., y Mark, M. M. (2003b). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, 24 (3), 293-312.
- Herman, R.E y Gioia, J.L. (1998). *Lean and Meaningful*. Winchester VA. Oakhills Press.
- Hopson, R.K. (2000). 'Editor's Notes'. En R. K. Hopson (Ed), *How and why language matters in evaluation*. New directions for evaluation, 86. San Francisco. Jossey-Bass.
- Horton, D, et al (2003). Evaluating Capacity Development. *Experiences from Research and Development Organizations around the World*. IDCR.
- House, E. (1993). *Professional evaluation: Social impact and political consequences*. Newbury Park, CA. Sage.
- Huber, G. (1991). Organizational learning: the contributing processes and the literatures. *Organization Science*, 2 (1), 88-115.
- Huffman, D., Lawrenz, F., Thomas, K., y Clarkson, L. (2006). Collaborative evaluation communities in urban schools: A model of evaluation capacity building for STEM education. *New Directions for Evaluation*, 109, 73-85

- Huffman, D., Thomas, K., y Lawrenz, F. (2008). A Collaborative Immersion Approach to Evaluation Capacity Building. *American Journal of Evaluation*, 29, 358-368
- Hunter, D. E. K. (2006). Using a theory of change approach to build organizational strength, capacity and sustainability with-not-for profit organizations in the human services sector. *Evaluation and Program Planning*, 29 (2) 193-200.
- Jay, M, O'Sullivan, R. G., and Costello, E. P. (2006). *Using Collaborative Evaluation to Capture GEAR UP North Carolina's Programme Outcomes*. Presentado en la conferencia anual de la Asociación Americana de Investigación Educativa. San Francisco, CA, April 11, 2006.
- Jenlink, P.M. (1994). Using evaluation to understand the learning architecture of an organization. *Evaluation and Program Planning*, 17 (3), 315-325.
- Johnson, K., Greenseid L., Toal, S., King, J., Lawrenz, F. y Volkov, B. (2009). Research on Evaluation Use: A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, 30, 377-410.
- Joint Committee on Standards for educational evaluation (1981). *Standards for evaluations of educational programs, projects and materials*, New York. Mc Graw-Hill.
- Joint Committee on Standards for educational evaluation (1994). *The program evaluation standards*. Thousand Oaks. CA. Sage.
- Joint Committee on Standards for educational evaluation (2011). *The program evaluation standards*. Newbury Park. CA. Sage.
- Joint Inspection Unit (2014). *Analysis of the Evaluation Function in the UN System*. JIU/REP/2014/6.
- Kaplan, S.A., y Garrett, K.E. (2005). The use of logic models by community-based initiatives. *Evaluation and programme planning*, 28, 167-192
- Kassam, Y., y Mustafa, K. (1982). *Participatory Research: An Emerging Alternative Methodology in Social Science Research*. Toronto: International Council for Adult Education.
- Kellogg Foundation, W. K. (2001). *Logic model development guide*. Extraído de: <http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf>
- Kettner, P., Moroney, R., y Martin, L. (1999). *Designing and managing programs: An effectiveness-based approach*. Thousand Oaks, CA: Sage
- King, J. (1998). Making sense of participatory evaluation practice. *New Directions for Evaluation*, 80, 57-67.
- King, J.A. (1995). Research on evaluation and its implications for evaluation research and practice. *Studies in Educational Evaluation*, 14, 285-299.
- King, J.A. (2002). Building the evaluation capacity of a school district. En D.W. Compton, M. Baizerman y S.H. Stockdill (Eds), *The art, chat and sáciense of*

- evaluation capacity building*. New directions for Evaluation, 93, 63-80. San Francisco, CA: Jossey-Bass.
- King, J.A. (2007). Developing evaluation capacity through process use. *New Directions for Evaluation*, 126, 45-59.
  - King J. A., y Pechman, E.M (1984). Pinning the wave to the shore: Conceptualizing evaluation use in school systems. *Educational Evaluation and Policy Analysis*, 6, 241-253.
  - King, J. A., y Volkov, B. (2005). A framework for building evaluation capacity based on the experiences of three organizations. *CURA Reporter*, 35 (3), 10-16.
  - Kirkhart, K, E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. En V. Caracelli and H. Preskill (eds) *The Expanding Scope of Evaluation Use*, New Directions for Evaluation, 88, 5-23. San Francisco: Jossey-Bass.
  - Kotter, J. (1996). *Leading change*, Cambridge, MA: Harvard Business School Press.
  - Kuzmin, A. (2005). *Exploration of Factors That Affect the Use of Evaluation Training in Evaluation Capacity Development*. Union Institute and University, Cincinnati, Ohio.
  - Labin, S., Duffy, J., Meyers, D.C., Wandersman, A., y Lesesne, C.A (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, 33, 307-338.
  - Ledermann, S. (2012). Exploring the necessary conditions for evaluation use in program change. *American Journal of Evaluation*, 33 (2), 159-178.
  - Leeuw, F. (2006). Managing evaluations in the Netherlands and Types of Knowledge. En Rist, R., y Stame, N. (Eds) *From studies to streams: Managing evaluative systems*. Comparative policy evaluation.
  - Leithwood, K., y Montgomery, D. (1987). *Improving classroom practice using innovation profiles*. Toronto: OISE Press.
  - Leviton, L. (2003). Evaluation Use: Advances, challenges and applications, *American Journal of Evaluation*, 24, 525-535.
  - Leviton, L.C., y Hughes, E.F.X. (1981). Research on the utilization of evaluations: A review and a synthesis. *Evaluation Review*, 5 (4), 525-548.
  - Leviton, L. y Boruch R. F. (1983). Contributions of evaluation to education programs and policy. *Evaluation Review*, 7 (5), 563-598.
  - López, G., López, M. B. (2003). *Aprendizaje Cooperativo y Colaborativo. Su Implementación En Carreras Universitarias*, Actas del TICEC 05 297 Congreso Latinoamericano de Educación Superior del Siglo XXI, Vol II, 1-9.
  - Lord, J., y Huchison, P. (2003). Individualized support and funding: Building blocks for capacity building and inclusión. *Disability and Society*, 18 (1), 71.
  - Louis, K.S. (1994). Beyond “managed change”: Rethinking how schools improve. *School Effectiveness and School Improvement*, 5 (1), 2-24.

- Love, A. (2006). Developing evaluation capacity in extension 4-H field faculty: A framework for success. *American Journal of Evaluation*, 22, 257–269.
- Mackay, K. (2002). *Annual Report on Evaluation Capacity Development*. Washington D.C.: Operations Evaluation Department, The World Bank.
- Mackie, J.L. (1967). *Mill's Methods of Induction*. En P. Edwards, ed., *The Encyclopedia of Philosophy* (8 vol.), New York: MacMillan, vol. 5, 324-332.
- Mahoney, J., y Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis*, 14, 227-249.
- Mark, M.M., y Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10, 35-57.
- Mark, M. M., y Shotland, R. L. (1985). Stakeholder-Based Evaluation and Value judgments: The Role of Perceived Power and Legitimacy in the Selection of Stakeholder Groups. *Evaluation Review*, 9, 605-626.
- Mathison, S. (2005). *Encyclopedia of evaluation*. SAGE publications
- Mayne, J. (2001). Addressing attribution through contribution analysis: using performance measures sensibly. *Canadian Journal of Program Evaluation* 16, 1-24.
- Mayne, J. (2007). Challenges and lessons in implementing Results Based Management. *Evaluation* 13 (1), 87-109.
- Mayne, J, y Zapico-Goñi, E. (1997). *Monitoring performance in the public sector: Future directions from international experience*. New Brunswick NJ. Transaction Publishers.
- Mayne, J. y Rist, R.C. (2006). Studies are not enough: the necessary transformation of evaluation. *Canadian Journal of Program Evaluation*, 21, 93-120.
- Milstein, B., Chapel, T. J., Wetterhall, S. F., y Cotton, D. A. (2002). Building capacity for program evaluation at the Centers for Disease Control and Prevention. *New Directions for Evaluation*, 93, 27-46. San Francisco: Jossey-Bass.
- Miller, R., y Campbell, R. (2006). Taking stock of empowerment evaluation: An empirical review. *American Journal of Evaluation*, 27, 296–319
- Monterde Díaz, R. (2014). Calidad y Utilidad de las Evaluaciones en la Cooperación Española al Desarrollo. *Instituto Universitario de Estudios Europeos*. Documento de Trabajo Serie CECOD. Número 24/2014.
- Moxnes, P. (1981). *Anxiety and work environment*. Stockholm: Natur och Kultur.
- Mueller, M. R., y Fitzpatrick, J. L. (1998). The evaluation of Minnesota's Early Childhood Family Education Program. Dialogue with Marsha Mueller. *American Journal of Evaluation*, 19, 80–99.
- Nielsen, S. B., y Ejler, N. (2008). Improving performance? Exploring the complementarities between Evaluation and Performance Management. *Evaluation*. 14 (2), 171-192.

- Nielsen, S.B., Lemire, S., y Skov, M. (2011). Measuring evaluation capacity: Results and implications of a Danish study. *American Journal of Evaluation*, 32, 324-344.
- Niemiec, R. P. y Walberg, H. J. (1992). The effect of computers on learning. *International Journal of Education*, 17, 99-108.
- Nonaka, I., y Takeuchi, H. (1999). *La organización creadora de conocimiento*. Oxford University Press, México.
- Nye, N., y Glickman, N.J. (2000). Working together: Building capacity for community development. *Building Capacity for Community Development*, 11 (1), 163-198
- O'Sullivan, R. (2004). *Practicing evaluation: A collaborative approach*. Thousand Oaks, CA: Sage.
- O' Sullivan R. G (2007). Collaborative evaluation. Creating environments to engage stakeholders in evaluation. En *From Policies to Results: Developing Capacities for Country Monitoring and Evaluation Systems*. UNICEF.
- O'Sullivan, R. G., and D'Agostino, A. (2002). Promoting evaluation through collaboration: Findings from community-based programmes for young children and their families. *Evaluation*, 8 (3), 372-387.
- O'Sullivan, R. G., Skaga, J., and Chernow, M. (2005). First 5 Los Angeles Child Care Initiative Evaluation Phase II Final Report. Chapel Hill, NC: *Evaluation, Assessment and Policy Connections*, School of Education, University of North Carolina at Chapel Hill.
- OECD (1991). *Principles for Evaluation of development assistance*.
- OECD (1997). *Search for Results*, Public Management Practices
- OECD (1998). *PUMA Policy Brief*, Best practice guidelines for Evaluation,5.
- OECD (2002). *Glossary of Evaluation and Results Based Management (RBM) Terms*.
- OECD (2005). Aid Effectiveness: Three Good Reasons Why the Paris Declaration Will Make a Difference. En *2005 Development Co-Operation Report*.
- OECD (2006). *DAC in Dates. The History of OECD's*. Development Assistance Committee. OCDE Publications.
- OECD (2009). *Supporting evaluation capacity development in partner countries*. Room Document IV.
- OECD (2010). *Effective practices in conducting a multi-donor evaluation*. Evaluation and Aid Effectiveness.
- OIOS (2007a). *Assessment of evaluation capacities and needs in the United Nations Secretariat*, IED-2006-006, 24 August 2007.
- OIOS (2007b). *Guidance to programmes for developing an evaluation policy*.
- OIOS (2015). *Strengthening the role of evaluation and the application of evaluation findings on programme design, delivery and policy directives*. A/70/72

- Owen J.M. (1993). *Program evaluation. Forms and approaches*. St Leonards, Australia: Allen and Unwin.
- Owen, J.M. (2003). Evaluation culture: A definition and analysis of its development within organizations. *Evaluation Journal of Australasia*, 3, 43–47.
- Owen, J.M. (2006). *Program Evaluation: Forms and Approaches*, 3rd Edition. Sydney: Allen and Unwin.
- Owen J.M y Lambert, F.C. (1995). Roles for evaluation in learning organizations. *Evaluation 1* (2), 237-250.
- Palacios-Maldonado, M. (2000). Aprendizaje organizacional: Conceptos, procesos y estrategias. *Hitos de ciencias económico administrativas 15*, 31-39.
- Palumbo, D.J (1994). The political roots of misuse of evaluation. En C. J. Stevens y M. Dial eds. *New directions for program evaluation: Preventing the misuse of evaluation*, 64, 15-24. San Francisco: Jossey Bass.
- Patton M. (1997). *Utilization Focused Evaluation*. The New Century Text. 3<sup>rd</sup> Edition. Thousand Oaks. CA: Sage.
- Patton, M. (1988a). The evaluator's responsibility for utilization. *Evaluation Practice 9* (2), 5-24.
- Patton, M. (1988b). Six honest serving men for evaluation. *Studies in educational evaluation*, 14, 301-330.
- Patton, M. (1998). Discovering Process use. *Evaluation 4* (2), 225-233.
- Patton, M. (2002). *Qualitative research and evaluation methods (3<sup>rd</sup> ed.)*. Thousand Oaks. CA. Sage.
- Patton, M. (2005). Misuse of evaluations. En 'Encyclopedia of evaluation', S. Mathiason. Thousand Oaks, CA: Sage.
- Patton, M. Q. (1996). A world larger than formative and summative. *Evaluation Practice*, 17 (2), 131–144.
- Patton, M. (2008). *Utilization Focused Evaluation 4<sup>rd</sup> Edition*. Sage.
- Patton, M. Q. (2007). Process use as a usefulness. En J. B. Cousins (Ed.), Process use in theory, research, and practice. *New Directions for Evaluation*, 116, 99–112.
- Patton, M. Q. (2012). *Essentials of utilization-focused evaluation (4<sup>th</sup> ed.)*. Thousand Oaks, CA: Sage.
- Pawson, R. y Tilley. (1997). *Realistic Evaluation*. London: Sage.
- Pennings, P. (2003). Beyond dichotomous explanations: explaining constitutional control of the executive with fuzzy-sets. *European Journal of Political Research 42* (4), 541–567.
- Pérez Liñán, A. (2007). *Instrucciones para utilizar fs/QCA*. Universidad de Salamanca.

- Perrin, B. (1999). Performance Measurement: Does the reality match the rhetoric? A rejoinder to Bernstein and Winston. *American Journal of Evaluation* 20 (1),101-111.
- Perrin, B. (2006). How evaluation can help make knowledge management real. En Rist, R., y Stame, N. (Eds) *From studies to streams: Managing evaluative systems*. Comparative policy evaluation.
- Perrin, B. (2007). Towards a new view of accountability. En M.-J. Bemelmans-Videc, J. Lonsdale, & B. Perrin (Eds.), *Making accountability work: Dilemmas for evaluation and for audit*. New Brunswick, NJ: Transaction Publishers.
- Preskill, H. (1994). Evaluation's role in enhancing organizational learning. *Evaluation and Programme planning*, 17(3), 291-297.
- Preskill H. (2007). Exploring effective strategies for facilitating evaluation capacity development. En *From Policies to Results: Developing Capacities for Country Monitoring and Evaluation Systems*. UNICEF.
- Preskill, H y Torres, R.T (1996). *From evaluation to evaluative inquiry for organizational learning*. Presentado en la conferencia anual de la AEA. Atlanta.
- Preskill, H., Caracelli, V.J. (1996). *The past, present and future conceptions of evaluation use: Results from a survey on current conceptions of evaluation use*. Presentado en la conferencia anual de la AEA. Atlanta.
- Preskill, H., y Caracelli, V. (1997). Current and developing conception of use: Evaluation using TIG survey results. *Evaluation Practice*, 18 (3), 209-226.
- Preskill, H. y Torres, R. T. (1999). *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, CA: Sage.
- Preskill, H., Zuckerman, B. y Matthews, B. (2003). An exploratory study of process use: findings and implications for further research. *American Journal of Evaluation* 24 (4), 423-442.
- Preskill, H. y Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29, 443-459.
- Programa de Acción de Accra <http://www.oecd.org/dataoecd/58/19/41202043.pdf>
- Ragin, C.C. (1987). *The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Ragin, C. C. (2000). *Fuzzy-Set Social Science*. Chicago and London: The University of Chicago Press.
- Ragin, C.C. (2006). Set relations in social research: evaluating their consistency and coverage. *Political Analysis* 14, 291–310.
- Ragin, C.C. (2008). *Redesigning Social Inquiry. Fuzzy Sets and Beyond*. Chicago and London: The University of Chicago Press.
- Ragin, C.C. (2009). Qualitative Comparative Analysis. Using fuzzy sets (fsQCA). En Rihoux and Ragin (eds.), pp. 87–121.

- Ragin, C. C. y Rihoux, B. (2004) 'Qualitative Comparative Analysis (QCA): State of the Art and Prospects', *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 2 (2), 3–13.
- Ragin, C.C. y Sonnett, J. (2005). Between complexity and parsimony: Limited diversity, counterfactual cases and comparative analysis. En S. Kropp y M. Minkenberg (eds), *Vergleichen in der Politikwissenschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Rebien, C. (1996). *Evaluating Development Assistance in Theory and Practice*. Aldershot: Avebury.
- Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation, ST/SGB/2000/8, 19 April 2000.
- Rich, F. (1977). Uses of social science information by federal bureaucrats: knowledge for action versus knowledge for understanding. En C. H. Weiss (ed.). *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books
- Rihoux, B, y Grimm, H. (2006). *Innovative Comparative Methods for Policy Analysis. Beyond the Quantitative–Qualitative Divide*. New York: Springer.
- Rist, R. (2006). The 'E' in Monitoring and Evaluation: Using evaluative knowledge to support a Results Based Management System. En Rist, R., y Stame, N. (Eds) *From studies to streams: Managing evaluative systems*. Comparative policy evaluation.
- Rist, R. y Morras, L. (2004). *The road to results*, The world bank.
- Ristau, S. (2001). Building organizational capacity in outcomes evaluation: A successful state association model. *Families in society*, 82 (6), 555-560.
- Robbins, S.P. (1998). *Organisational behaviour. concepts, controversies, and applications (8th ed.)*, Upper Saddle River, New Jersey: Prentice Hall.
- Rodríguez-Ariza, C. (2014). La gestión del conocimiento y el uso de las evaluaciones. Teoría, retórica y práctica. *Instituto Universitario de Estudios Europeos*. Documento de Trabajo Serie CECOD. Número 27/2014.
- Rodríguez-Campos, L. (2005). *Collaborative evaluations*. Tamarac, FL: Llumina Press.
- Rodríguez-Campos, L., y Rincones-Gómez, R. (2013). *Collaborative evaluations: Step-by-step (2nd ed.)*. Stanford, CA: Stanford University Press.
- Rossi, P.H y G. F. Freeman. (1985). *A systematic approach. 3<sup>rd</sup> Edition*. Beverly hills, California, SAGE.
- Rossi, P.H., Lipsey M.W., y Freeman, H.E. (2004). *Evaluation: A systematic approach, 7<sup>th</sup> Edition*, SAGE publications.
- Rowe, W.E., y Jacobs, N.F. (1998). Principles and practices of organizationally integrated evaluation. *Canadian Journal of Program Evaluation*, 13(1), 115–138.
- Ryan, K., y Johnson, T. (2000). Democratizing evaluation: Meanings and methods from practice. En K. Ryan y L. DeStefano (Eds.), *Evaluation as a democratic*

- process: Inclusion, dialogue and deliberation. *New directions for evaluation*, 85, 39–55. San Francisco, CA: Jossey Bass.
- Sager, F (2005). Metropolitan Institutions and Policy Coordination. The Integration of Land Use and Transport. *Policies in Swiss Urban Areas, Governance* 18 (2), 227-256.
  - Sanahuja, J.A. (2007). ¿Más y mejor ayuda? La Declaración de París y las tendencias en la cooperación al desarrollo, en Manuela Mesa (Coord.), *Guerra y conflictos en el Siglo XXI: Tendencias globales*. Anuario 2007-2008.
  - Sanders, J. R. (2002). Presidential address: On mainstreaming evaluation. *American Journal of Evaluation*, 23 (3), 253-259.
  - Sanders, J. R. (2003). Mainstreaming evaluation. *New Directions for Evaluation*, 99, 3–6.
  - Sanders, J.R (2004). *The Program Evaluation Standards*. Joint Committee on Standards for Educational Evaluation. London. Sage
  - Sanders, J.R (1995). Standards and principles. *New Directions for Program Evaluation*, 66, 47-52. San Francisco: Jossey-Bass.
  - Saunders, M. (2012). The use and usability of evaluation outputs. *Evaluation*, 18 (4), 421-436.
  - Scriven, M. (1967). The methodology of Evaluation. En R. Tyler, R. Gagné y M. Scriven (Eds.), *Perspectives of Curriculum Evaluation* (AERA Monograph Series on Curriculum Evaluation, No. 1) (Chicago, Rand McNally), 39-83.
  - Scriven M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17(3), 151-161.
  - Scriven, M. (1991). *Evaluation thesaurus (4th ed.)*. Newbury Park, CA: Sage.
  - Scriven, M. (1994a). Product evaluation--The state of the art. *American Journal of Evaluation*, 15 (1), 45-62.
  - Scriven, M. (1994b). The final synthesis. *American Journal of Evaluation*, 15 (3), 367-382.
  - Schaumburg-Muller, H. (1996) *Support to Evaluation Capacity Building: Donor Experiences and Lessons Learned*. DAC Expert Group on Aid Evaluation, Canberra, Australia, 6–8 February.
  - Scheirer, M.A. (1987). Program theory and implementation theory: implications for evaluators, *New directions for evaluation*, 33, 59-76.
  - Schneider, C.Q. y Wagemann, C. (2007). *Qualitative Comparative Analysis (QCA) und Fuzzy Sets. Ein Lehrbuch für Anwender und jene, die es werden wollen*. Opladen/Farmington Hills, MI: Barbara Budrich.
  - Schneider, C. Q. y Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: A Guide for Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.

- Sechrest, L. (1994). Program evaluation: Oh what it seemed to be! *Evaluation Practice*, 15 (3), 359-367.
- Segone, M. (2007). Moving from Policies to Results by Developing National Capacities for Country-Led Monitoring and Evaluation Systems. En *From Policies to Results: Developing Capacities for Country Monitoring and Evaluation Systems*. UNICEF.
- Senge, P. (1990). *The fifth discipline: The Art and Practice of the Learning Organization*. Century Business, London.
- Shadish, W.R. (1987). Program micro-and macrotheories: a guide for social change. *New directions for evaluation*, 33, 93-109.
- Shadish, W. R. (1994). Need-based evaluation theory: What do you need to know to do good evaluation? *Evaluation Practice*, 1S, 347-358.
- Shadish, W.R. y Epstein R.A (1987). Patterns of program evaluation practice among members of the Evaluation Research Society and Evaluation Network. *Evaluation Review*, 11(5), 555-590.
- Shadish W. y Luellen, J. (2005). History of evaluation, pp. 183-86. En *Encyclopedia of Evaluation*, S. Mathison. Thousand Oaks, CA. Sage.
- Shulha, L y Cousins B. (1997). Evaluation use: Theory, research and practice since 1986. *American Evaluation Association*, 18, 195-208.
- Simon H. A. (1991). Bounded rationality and organizational learning. *Organization Science* 2(1), 125-134.
- Smith, N. L. (2007). Empowerment evaluation as evaluation ideology. *American Journal of Evaluation*, 28 (2), 169-178.
- Sobeck, J., y Agius, E. (2007). Organizational capacity building: Addressing a research and practice gap. *Evaluation and Programme planning*, 30, 237-246.
- Sociedad Alemana de Evaluación: *Evaluation Standards*.
- Sociedad Australasiática de Evaluación: *Guidelines on Ethical Conduct of Evaluation*.
- Sociedad Británica de Evaluación: *Good Practice Guidelines*.
- Sociedad Canadiense de Evaluación: *Guidelines for Ethical Conduct*.
- Sociedad Francesa de Evaluación: *Carta de la Evaluación de las Políticas Públicas y de los Programas Públicos*.
- Sonnichsen, R. (1999). *High impact internal evaluation: A practitioner's guide to evaluating and consulting inside organizations*. Thousand Oaks, CA: Sage.
- Stadler Waerness, M.S., Ovrelid R. (2006). Evaluative information in Norwegian Ministries. En Rist, R., y, N. (Eds) *From studies to streams: Managing evaluative systems*. Comparative policy evaluation.
- Stahler, G.J. (1995). Improving the quality of federal human services national demonstration programs. *Evaluation and Program Planning*, 18 (2), 129-141.

- Stevens, C.J, Dial M. (1994). What constitutes misuse? En Stevens and Dial Eds. *New Directions for program evaluation: Guiding principles for evaluators*, 64 (pp 3-14). San Francisco: Jossey-Bass.
- Stevenson, J. F., Florin, P., Mills, D. S., y Andrade, M. (2002). Building evaluation capacity in human service organizations: A case study. *Evaluation and Program Planning*, 25, 233–243.
- Stockdill, S.H., Baizerman, M. y Compton, D.W (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, 93, 7-25.
- Stufflebeam, D.L. (1987). *Evaluación sistemática. Guía teórica y práctica*. Madrid: Paidós/MEC.
- Stufflebeam, D. (1994). Empowerment evaluation, objectivist evaluation and evaluation standards: where the future of evaluation should not go, and where it needs to go. *Evaluation Practice*, 15(3) 321-338.
- Stufflebeam, D., y Shinkfield, A.J. (1985). *Systematic Evaluation: A self-instructional guide to theory and practice*. Norwell: MA, Kluwer.
- Stufflebeam, D., y Shinkfield, A.J. (2007). *Evaluation Theory, Models, and Applications*. Jossey-Bass, San Francisco.
- Suarez-Balcázar, Y. (2002). Aspectos de diversidad humana en la evaluación de programas comunitarios. *Intervención Psicosocial*, 11, 167-182.
- Suarez-Balcázar, Y., Taylor-Ritzler, T., García-Iriarte, E., Keys, C. B., Kinney, L., y Ruch-Ross, H., Curtin, G. (2010). Evaluation capacity building: A culturally- and contextually-grounded interactive framework and exemplar. En F. Balcazar, Y. Suarez-Balcazar, T. Taylor-Ritzler, y C. B. Keys (Eds.), *Race, culture and disability: Rehabilitation science and practice* (pp. 307–324). Sudbury, MA: Jones and Bartlett.
- Sutherland, S. (2004). Creating a culture of data use for continuous improvement: A case study of an Edison project school. *American Journal of Evaluation*, 25, 277–293.
- Tamayo Sáez, M. (1997) El análisis de las políticas públicas. En Bañón, R. y Carrillo, E. (comps.) *La Nueva Administración Pública*. Madrid: Alianza.
- Tandon, R. (1981). Participatory research in the empowerment of people. *Convergence*, 14, 20–29.
- Tandon, R. (2008). Participation, citizenship and democracy: Reflections on 25 years of PRIA. *Community Development Journal*, 43, 284–296.
- Taut, S. (2007). Studying evaluation capacity building in a large international development organization. *American Journal of Evaluation*, 28 (1), 45-59.
- Taylor-Ritzler, T., Suarez-Balcazar, Y., García-Iriarte, E., Henry, D., y Balcazar, F. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation*, 34, 190–206.

- Thoenig, J.C. (2000). Evaluation as Usable knowledge for public management reforms. *Evaluation*, 6 (2), 217-229.
- Torres, R.T, Preskill H.S y Piontek, M.E (1996). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Thousand Oaks, CA. Sage.
- Torres, R. T., y Preskill, H. (2001). Evaluation and organizational learning: Past, present and future. *American Journal of Evaluation*, 22 (3), 387–395.
- Toulemonde, J. (1999). Incentives, constraints and culture-building as instruments for the development of evaluation demand. En R. Boyle y D. Lemaire (Eds.), *Building effective evaluation capacity* (pp. 153-174). New Brunswick, NJ: Transaction Publishing.
- Trochim, W. (2006). Face validity. Retrieved from the Center for Social Research Methods Web Site: [http://www .socialresearchmethods.net/kb/measval.htm](http://www.socialresearchmethods.net/kb/measval.htm)
- UN Habitat (2012). *Evaluation brief. Evaluation of the Implementation of UN-Habitat's Medium-Term Strategic and Institutional Plan 2008-2013*. Evaluation Report 1/2012.
- UNDP (2008). *Supporting Capacity Development. The UNDP approach*. Bureau for development policy. New York.
- UNEG (2005). *Norms and Standards*.
- UNEG (2007). *Evaluation in the UN System: UNEG working group on Evaluation and Oversight*.
- UNEG/FAO (2012). *Peer Review of the Evaluation Function of FAO*. <http://www.uneval.org/document/detail/1191>
- UNEG/GEF (2009). *Peer Review of the Evaluation Function of GEF*. <http://www.uneval.org/document/detail/313>
- UNEG/GEF (2014). *Second Peer Review of the Evaluation Function of GEF*. <http://www.uneval.org/document/detail/1624>
- UNEG/OIOS (2009). *Peer Review of the Evaluation Funcion of OIOS*. <http://www.uneval.org/document/detail/226>
- UNEG/PMA (2007). *Peer Review of the Evaluation Funcion of PMA*. <http://www.uneval.org/document/detail/97>
- UNEG/PMA (2014). *Second Peer Review of the Evaluation Function of PMA*. <http://www.uneval.org/document/detail/1623>
- UNEG/REF/2011/1. *UNEG Framework for professional peer reviews of the evaluation function of UN organizations*. <http://www.uneval.org/document/detail/945>
- UNEG/UN Habitat (2012). *Peer Review of the Evaluation Funcion of UN Habitat*. <http://www.uneval.org/document/detail/1164>
- UNEG/UN Women (2014). *Peer Review of the Evaluation Function of UN Women*. <http://www.uneval.org/document/detail/1626>

- UNEG/UNDP (2006). *Peer Review of the Evaluation Function of UNDP*. <http://www.uneval.org/document/detail/99>
- UNEG/UNDP (2013). *Second Peer Review of the Evaluation Function of UNDP*. <http://www.uneval.org/document/detail/1632>
- UNEG/UNEP (2012). *Peer Review of the Evaluation Function of UNEP*. <http://www.uneval.org/document/detail/1163>
- UNEG/UNICEF (2006). *Peer Review of the Evaluation Function of UNICEF*. <http://www.uneval.org/document/detail/98>
- UNEG/UNIDO (2010). *Peer Review of the Evaluation Function of UNIDO*. <http://www.uneval.org/document/detail/655>
- UNICEF Universalia (2013). *Review of the Evaluation Function*. [http://www.unicef.org/evaluation/files/UNICEF\\_Synthesis\\_Report\\_Revised\\_Final.pdf](http://www.unicef.org/evaluation/files/UNICEF_Synthesis_Report_Revised_Final.pdf)
- United Way of America. (1996). *Measuring program outcomes: A practical approach*. Arlington, VA: Author.
- United Way of America. (2008). *Logic model handbook*. Arlington, VA: Author. [http://www.vsuw.org/file/logic\\_model\\_handbook\\_updated\\_2008.pdf](http://www.vsuw.org/file/logic_model_handbook_updated_2008.pdf)
- Van Thiel, S. y Leeuw, F.L (2002). The Performance Paradox in the Public Sector. *Public Performance and Management Review* 25 (3), 267-281.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. Transaction Pub.
- Verkuilen, J. (2005). Assigning membership in a fuzzy set analysis. *Sociological Methods and Research* 33 (4), 462–496.
- Volkov, B. B., y King, J. A. (2007). *A checklist for building organizational evaluation capacity*. <http://www.wmich.edu/evalctr/checklists/ecb.pdf>
- Vroom, P.I, Colombo M. y Nahan, N. (1994). Confronting ideology and self-interest: Avoiding misuse of evaluation. En C.J Stevens y M. Dial Eds. *New directions for program evaluation: Preventing the misuse of evaluation*, 64 (pp 49-60). San Francisco: Jossey Bass.
- Wandersman, A., Imm, P., Chinman, M., & Kaftarian, S. (2000). Getting to outcomes: Methods and tools for planning, evaluation and accountability. Rockville, MD: Center for Substance Abuse Prevention.
- Weaver, L., y Cousins, J. B. (2004). Unpacking the participatory process. *Journal of Multidisciplinary Evaluation*, 1, 19-40.
- Weiss, C. (1980). Knowledge Creep and Decision Accretion'. Knowledge: Creation, Utilization, Diffusion. *Science Communication*, 1 (3), 381-404.
- Weiss, C. (1988a). *Evaluation. 2<sup>nd</sup> Edition*. Prentice Hall. New Jersey.
- Weiss C. (1988b). Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice*, 9(1), 5-19.

- Weiss C. (1988c). If program decisions hinged only on information: A response to Patton. *Evaluation Practice*, 9(3), 15-28.
- Weiss, C. (1998). Have we learned anything new about the use of evaluation? *AEA* 19, 21-34.
- Weiss, C. (1999). The interface between Evaluation and Public Policy. *Evaluation*, 5, 468- 486.
- Weiss, C. (2004). Rooting for Evaluation: A cliff notes version of my work- p 153-168. *Evaluation Roots: Tracing Theorists, Views and Influences*, Alkin. Sage.
- Whitney, D., Trosten-Bloom, A., and Cooperrider, D. (2003). *The power of Appreciative Inquiry: A practical guide to positive change*. San Francisco: Berrett-Koehler.
- Worthen, B. R. (1999). Critical challenges confronting certification of evaluators. *American Journal of Evaluation*, 20, 533-555.
- Yin, R., and White, J. (1984). *Federal Technical Assistance Efforts: Lessons and Improvement in Education for 1984 and Beyond*. Washington, DC: Cosmos Corporation.
- Young, G. (2006). Evaluation can cross the boundaries: The case of transport Canada. *Canadian Journal of Evaluation* 21 (3), 73-92.



## Anexos



## **Anexo 1: Las entidades objeto de estudio**

### **1.1 UNO Habitat**

#### Mandato

UNO Habitat, o también conocido como Programa de las Naciones Unidas para los Asentamientos Humanos, tiene como misión principal la promoción de la urbanización sostenible y reducción de la pobreza. Su establecimiento es relativamente reciente. De hecho, a pesar de que desde su creación la Asamblea General del sistema de NU había realizado algunos llamamientos incipientes para la prevención de los riesgos inherentes al crecimiento rápido y no siempre controlado de las áreas urbanas, no fue hasta la década de los 70, cuando se tomaron acciones concretas.

Así, en 1975 la Asamblea General de NU crea la Fundación para los Asentamientos Humanos y UN Habitat (UNHHSF), en lo que se constituye como la primera entidad oficial de las NU para la urbanización. Su principal misión era la de apoyar a los programas nacionales sobre asentamientos humanos a través de capacitación técnica, particularmente a los países en vías de desarrollo.

En los 70, se celebraron dos conferencias internacionales en el marco del sistema de NU donde se reconocieron los desafíos inherentes a la urbanización y sus efectos sobre el medio ambiente. La primera de ellas fue la Conferencia de las Naciones Unidas sobre el Medio Humano, celebrada en Estocolmo en 1972; la segunda fue la Conferencia de las Naciones Unidas sobre Asentamientos Humanos, celebrada en Vancouver en 1976 (Habitat I). Esta segunda conferencia propició la creación en 1977 de la Comisión de Naciones Unidas para los Asentamientos Humanos (organismo intergubernamental), y el Centro de las Naciones Unidas para los Asentamientos Humanos (CNUAH), como secretaría ejecutiva de dicha Comisión. A partir de entonces, el CNUAH, fue también el encargado de gestionar los fondos de la Fundación UNHHSF.

Desde sus comienzos, y cerca lo largo de cerca de dos décadas, el CNUAH, desarrolló un trabajo considerado altamente técnico a pesar de su reducido tamaño, especialmente en los países en vías de desarrollo. En 1996, la Segunda Conferencia de Naciones Unidas sobre Humanidad Ambiental (Habitat II), adoptó la declaración de Estambul y la Agenda Habitat. Esta declaración fue adoptada por 171 países, e incluía cerca de 100 compromisos y 600 recomendaciones<sup>105</sup>.

Como resultado de esta Segunda Conferencia, el marco de acción de CNUAH se vio altamente fortalecido, puesto que se le encomendó el apoyo y seguimiento a la implementación de la Agenda Habitat. En 2002, a través de la resolución de la Asamblea General A56/206, el CNUAH pasó a ser el Programa de las Naciones Unidas

---

<sup>105</sup> <http://www.un.org/ga/Istanbul+5/declaration.htm>

para los Asentamientos Humanos, ONU Habitat. Su relevancia estratégica no ha dejado de crecer desde entonces como parte integrante de las prioridades y agenda de desarrollo de NU. UNO Habitat es la responsable de monitorear e informar del progreso de las metas sobre el acceso a agua potable, y el acceso a sanidad básica en el marco de los Objetivos de Desarrollo del Milenio para el 2015.

Como se ha mencionado, la adopción de la Agenda Habitat en 1996 llevó a un rápido crecimiento en las actividades, personal y presupuesto de UNO Habitat. El programa fue evaluado por la Oficina de Supervisión Interna (OIOS) en 2005, y entre sus recomendaciones se incluyó la necesidad de priorizar sus actividades, para poder conseguir un mayor impacto.

En consecuencia al informe de OIOS en 2005, la Asamblea General solicitó a UNO Habitat la realización de un Plan Estratégico e Institucional de Medio Plazo (*Medium Term Strategic and Institutional Plan, MTSIP*). Con la puesta en marcha de este plan se pretendía delimitar estratégicamente el trabajo de UNO Habitat, darle una mayor coherencia a sus objetivos, y aplicar la gestión por resultados, y con ello, una mayor rendición de cuentas y transparencia en la gestión.

### Marco estratégico

El primer Plan Estratégico (MTSIP) para el periodo 2008-2013 fue adoptado por la resolución 21/2 de Abril de 2007 del Consejo de Gobierno (*Governing Council*). El plan estaba dividido en seis áreas específicas:

- ✓ Apoyo (*advocacy*), monitoreo, y colaboraciones (*partnership*)
- ✓ Planificación urbana participativa, gestión y gobernabilidad
- ✓ Acceso a tierra y a vivienda para todos/as
- ✓ Infraestructura y servicios básicos medioambientalmente sostenibles
- ✓ Sistemas financieros de asentamientos humanos
- ✓ Excelencia en gestión

Esta última área, excelencia en gestión, estaba enfocada en la puesta en marcha de un contexto facilitador para las otras cinco áreas sustantivas, a través de la gestión por resultados, la gestión del conocimiento, monitoreo y evaluación.

El MTSIP estaba diseñado para ser implementado a través de ciclos de programas de trabajo bianuales (2008-2009, 2010-2011, y 2012-2013). Los recursos necesarios para la implementación de las actividades englobadas en los tres programas de trabajo bianuales ascendieron a 319 Millones US\$ para el bienio 2008-2009, 357 Millones US\$ para el bienio 2010-2011, y 393 Millones US\$ para el bienio 2012-2013 (UNEG 2012UN Habitat 2012).

El MTSIP (2008-2013) fue evaluado en 2012. Las recomendaciones de la evaluación (UN Habitat 2012) llevaron a una reorganización de los objetivos estratégicos para mejorar en efectividad, eficiencia y optimizar el uso de recursos, y una reforma interna

de la organización<sup>106</sup>. En la actualidad está en vigor el segundo MTSIP, que cubre el periodo (2014-2019), y se compone de siete áreas temáticas de importancia estratégica:

- ✓ Marco regulatorio urbanístico, propiedad y gobernabilidad
- ✓ Planificación y diseño urbanístico
- ✓ Economía urbana
- ✓ Servicios básicos urbanos
- ✓ Viviendas y mejora de habitabilidad
- ✓ Reducción de riesgos y rehabilitación
- ✓ Investigación y desarrollo de capacidades

Entre dichas áreas, las primeras cuatro áreas temáticas tienen especial prioridad, puesto que se considera que el establecimiento de políticas adecuadas y un marco regulatorio urbanístico es clave para la consecución de otros objetivos de desarrollo económico y social. Así mismo, la planificación y diseño urbanístico es clave para la efectividad en el logro de resultados más relacionados con la implementación y operatividad de otras medidas para la reducción de riesgos y rehabilitación. El presupuesto para el bienio 2014-2015 ascendió a 336 Millones US\$, lo que supone una reducción del 17% con respecto al bienio anterior. Esta reducción forma parte de los esfuerzos por optimizar los recursos como consecuencia de la crisis económica mundial<sup>107</sup>.

### Gobernabilidad

ONU Habitat se compone de tres órganos fundamentales:

- ✓ El Consejo de Gobierno (GC), que establece las líneas estratégicas de la organización, y aprueba su programa y presupuesto. Está compuesto por 58 Estados Miembros, que representan 5 grupos regionales, elegidos por el Consejo Económico y Social de la Asamblea General de NNUU (ECOSOC). El Consejo de Gobierno se reúne cada 2 años.
- ✓ El Comité de Representantes Permanentes (CPR), que actúa como órgano de control y supervisión para el cumplimiento de las directrices establecidas por el Consejo de Gobierno. Representa al Consejo de Gobierno en todos aquellos asuntos pertinentes entre sus reuniones bianuales. El CPR está compuesto por embajadores o miembros del cuerpo diplomático de los Estados Miembros. Se reúne al menos en cuatro sesiones regulares al año, y se asiste de grupos de trabajo organizados en función de los temas a tratar.
- ✓ La Secretaría de ONU Habitat, que se encarga de ejecutar las decisiones del Consejo de Gobierno y hacerlas operativas a través de estrategias, programas y proyectos. Su sede está en Kenia, y es el punto focal de todos los asuntos

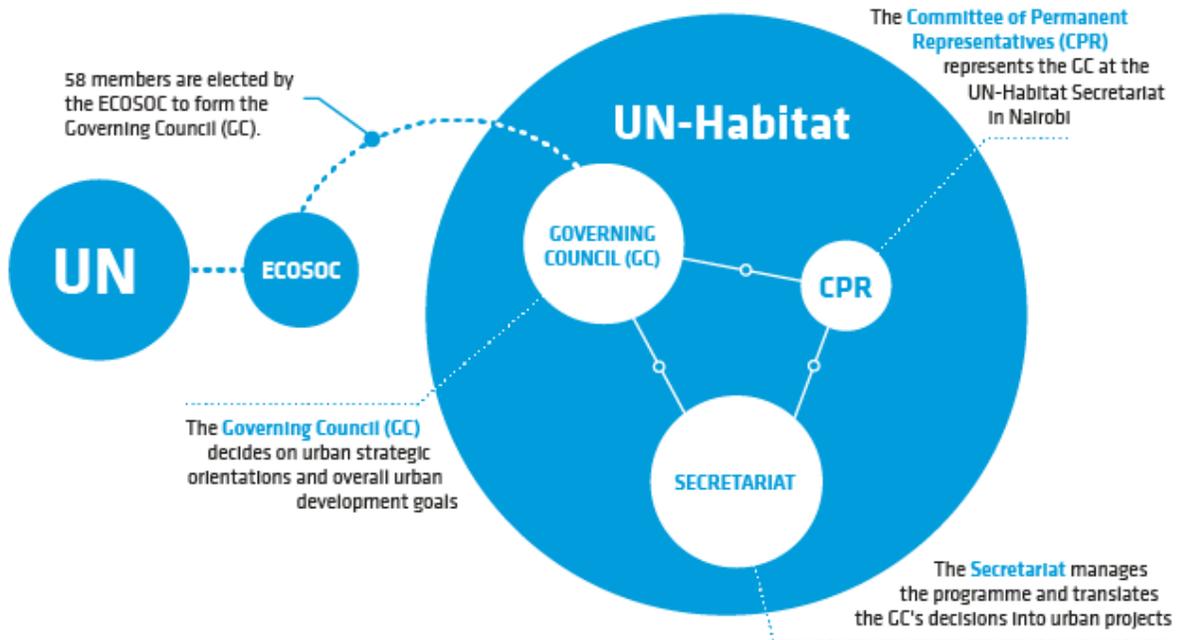
---

<sup>106</sup> Ref: Evaluation Report 1/2012, July 2012, UN Habitat

<sup>107</sup> Ref: HSP/GC/24/5 Proyecto de Programa y presupuesto para el bienio 2014-2015

relacionados con temas de urbanismo y asentamientos humanos en el marco de NU. La Secretaría se compone de: i) Oficinas, encargadas de áreas de gestión y apoyo; ii) Unidades temáticas, responsables de suministrar conocimiento técnico sobre un tema, y iii) las Oficinas regionales, que ejecutan el programa de trabajo a nivel regional o de país. El Director Ejecutivo es el máximo responsable de la Secretaría.

**Grafico A.1.1: Estructura organizacional UNO Habitat**



Fuente: <http://unhabitat.org/structure-overview/>

### Recursos

La fuente de fondos mayoritaria de ONU Habitat son las contribuciones voluntarias gubernamentales e intergubernamentales (cerca del 90%). En algunos casos, son contribuciones que se realizan para apoyar la implementación del programa de trabajo en general (20%), y otros casos, las contribuciones van a financiar proyectos específicos del plan del trabajo, y el programa de cooperación técnica (70%). Además de las contribuciones voluntarias, también existen las contribuciones regulares, las cuales son aprobadas por la Asamblea General de NU y forman parte del presupuesto 'troncal' de la organización (10%)<sup>108</sup>.

<sup>108</sup> UNEG/UN Habitat 2012

## La función de evaluación

La primera unidad de evaluación fue establecida en 1997 cuando todavía la organización era CNUAH. Sin embargo, la Unidad de Evaluación Independiente del Programa ONU Habitat fue establecida en el 2012, como parte del proceso de re-estructuración interna. La unidad tiene la responsabilidad de velar por el buen funcionamiento de la función de evaluación, y reporta directamente al Director Ejecutivo. La primera política de evaluación de UNO Habitat fue aprobada a nivel interno en 2013. La actual política de evaluación marca las directrices para la planificación y gestión de las evaluaciones de UNO Habitat.

La unidad de evaluación desarrolla un plan bianual de evaluaciones tanto las gestionadas directamente por la propia unidad de evaluación, como las gestionadas por los oficiales de programa de las unidades técnicas (también llamadas evaluaciones descentralizadas). La unidad de evaluación, también lleva a cabo otras funciones, como, por ejemplo, proporcionar capacitación sobre evaluación, desarrollar guías y procedimientos evaluativos, y llevar a cabo el seguimiento a las recomendaciones de los informes de evaluación<sup>109</sup>.

## El proceso de revisión por pares

La revisión por pares fue llevada a cabo entre Octubre 2011 y Enero 2012, en pleno proceso de revisión del primer MTSIP, por un panel multidisciplinar de evaluadores/as provenientes de otras agencias de NU (ONUDI, PNUD), y de organismos bilaterales de cooperación.

El panel efectuó una revisión exhaustiva documental de la institución, su estructura, procesos relacionados con la planificación y realización de evaluaciones. A su vez, seleccionó una muestra de informes de evaluación coordinados por la unidad de evaluación independiente para revisar su calidad en base a estándares UNEG.

Como parte de la metodología empleada, el panel también realizó entrevistas en persona y telefónicas a miembros de los órganos de gobierno (GC y CPR), personal ejecutivo de la Secretaría, personal de la unidad de evaluación, y responsables del resto de unidades técnicas. Además de las entrevistas, se distribuyó una encuesta por internet para poder acceder al personal que trabaja en oficinas regionales, de país y fuera de Nairobi.

El informe de revisión por pares fue distribuido y revisado por UNO Habitat, para su validación y comentarios. Su estructura incluye una descripción de la función de evaluación, la valoración del panel en torno a los criterios de utilidad, independencia y credibilidad, y una parte final donde se incluyen conclusiones y recomendaciones.

---

<sup>109</sup> UNEG/UN Habitat 2012 y <http://unhabitat.org/urban-knowledge/evaluation/>

El informe es públicamente accesible a través de la página web de UNEG (UNEG/UN Habitat 2012):

<http://www.uneval.org/document/detail/1164>

## 1.2 PNUD

### Mandato

El PNUD fue establecido como consecuencia de la fusión del Programa Ampliado de las Naciones Unidas para la Asistencia Técnica (creado en 1949), y el Fondo Especial de las Naciones Unidas (instituido en 1958). La Asamblea General del sistema de NU establece oficialmente el PNUD en 1965.

El mandato del PNUD va ligado a respaldar los procesos que se llevan a cabo a nivel nacional para acelerar el progreso en materia de desarrollo humano, a través del fomento de capacidades. Esto se resume en el primer Informe de Desarrollo Humano de 1990, donde se enfatiza que el concepto de desarrollo humano va ligado a la posibilidad de ampliar las opciones de la gente, considerando como parte de estas opciones tanto la formación de la capacidad humana como el uso que hace la gente de esta capacidad (PNUD 1990)

Para ello, el PNUD aporta soporte técnico en el desarrollo de políticas de apoyo a la gobernabilidad democrática, prevención de crisis, recuperación del medio ambiente y desarrollo sostenible. El apoyo técnico al desarrollo de políticas se basa en los siguientes principios clave<sup>110</sup>:

- ✓ Control y liderazgo nacionales
- ✓ Vínculo a los resultados de desarrollo a fin de lograr una mejora real en las vidas de las personas y opciones/oportunidades de las que disponen
- ✓ Flexibilidad para responder a las prioridades nacionales
- ✓ Desarrollo y uso de las capacidades, mecanismos y sistemas nacionales
- ✓ Fortalecimiento de la ejecución nacional
- ✓ Eficiencia, rendición de cuentas, resultados y transparencia
- ✓ Procedimientos simplificados y coherentes con los procedimientos nacionales

Además del apoyo técnico en el desarrollo de políticas, el PNUD tiene un rol importante de coordinación en las NU, como consecuencia del proceso de reforma requerido por la Asamblea General en 1997 (A/51/950). Entre las funciones de coordinación, se destacan por su relevancia las siguientes:

- ✓ Gestión del sistema de Coordinadores Residentes de NU. El PNUD gestiona y financia el sistema de Coordinadores Residentes. Los Coordinadores Residentes son nombrados por el Secretario General y son responsables por los equipos de país de las NU. El sistema de Coordinadores Residentes de NU está conformado por las diferentes entidades de NU, con independencia de su presencia formal en el país<sup>111</sup>.

---

<sup>110</sup>[http://www.undp.org/content/undp/es/home/operations/about\\_us.html](http://www.undp.org/content/undp/es/home/operations/about_us.html)

<sup>111</sup><https://undg.org/home/resident-coordinators/>

- ✓ Presidencia del Grupo de las Naciones Unidas para el Desarrollo (GNUM). El GNUM define las líneas estratégicas a seguir por el conjunto del sistema de NU en la coordinación, armonización, y coherencia de las iniciativas de desarrollo. El grupo está formado por los diferentes fondos, programas, organismos, departamentos y oficinas de NU<sup>112</sup>.
- ✓ Realización de los Informes sobre Desarrollo Humano. Los Informes sobre Desarrollo Humano son la publicación más importante del PNUD. Se publican desde 1990. Estos informes suelen tener un enfoque monográfico, e incluyen herramientas de medición, análisis y recomendaciones a políticas sectoriales. Los análisis de los informes se sustentan en el resultado de indicadores de desarrollo, entre los que se encuentran el Índice de Desarrollo Humano (IDH), el Índice de Desarrollo ajustado por Género, el Índice de Potenciación de Género y el Índice de Pobreza Humana<sup>113</sup>.
- ✓ Seguimiento a los Objetivos de Desarrollo del Milenio (ODM). El PNUD funciona como organismo de coordinación con los gobiernos nacionales para la realización de informes de progreso hacia los Objetivos de Desarrollo del Milenio a nivel de país. Los informes se realizan con el fin de reflejar los avances y facilitar un mayor enfoque en aquellos aspectos prioritarios y concretos para el desarrollo.<sup>114</sup>

### Marco estratégico

El PNUD formula planes estratégicos cuatrienales para establecer la dirección general de sus actividades. En el periodo 2008-2011, la acción del PNUD estaba organizada en torno a cuatro áreas específicas (DP/2007/43): 1) Logro de los ODM y reducción de la pobreza humana, 2) Fomento de la gobernabilidad democrática; 3) Apoyo a la prevención y recuperación de crisis y 4) Gestión de la energía y el medio ambiente para un desarrollo sostenible.

Para el periodo 2014-2017, las principales áreas de acción se concentran en tres esferas sustantivas de trabajo e interconectadas. La selección de estas esferas parte de un análisis de lo que el PNUD puede aportar en base a su mandato y valor añadido dentro del sistema de NU. *“El carácter institucional y los puntos fuertes básicos del PNUD configuran las decisiones que hemos adoptado respecto de nuestra labor futura. El PNUD no es un banco de desarrollo o una organización benéfica, ni es tampoco un organismo especializado ni una organización no gubernamental. Somos uno de los pocos organismos multilaterales de desarrollo que trabajan en forma holística en diversos sectores y en todo el mundo, operando en el núcleo de las actividades de*

<sup>112</sup><https://undg.org/home/about-undg/>

<sup>113</sup><http://hdr.undp.org/en/global-reports>

<sup>114</sup><http://www.undp.org/content/undp/en/home/librarypage/mdg/mdg-reports.html>

*desarrollo de las Naciones Unidas como administrador del sistema de coordinadores residentes y presidente del Grupo de las Naciones Unidas para el Desarrollo (GNUM)*” (DP/2013/40, pág. 6). Concretamente, las tres esferas de acción del actual plan cuatrienal son:

- ✓ Promocionar vías sostenibles de desarrollo. Para ello, el PNUD asiste a los países en la formulación y aplicación de vías de desarrollo que puedan hacer frente a la pobreza, la desigualdad y la exclusión a la vez que transformen la capacidad productiva. En consecuencia, se promueven todas aquellas estrategias orientadas a mejorar la dotación de recursos de los pobres e impulsar perspectivas de empleo y medios de vida.
- ✓ Reforzar la gobernabilidad democrática incluyente y efectiva. Esta esfera consiste en apoyar técnicamente a los países para que mantengan o logren una gobernabilidad pacífica y democrática en sus diferentes procesos de cambio, como procesos de reforma constitucional, la organización de elecciones creíbles o el refuerzo de los parlamentos.
- ✓ Aumentar la resiliencia<sup>115</sup>. Se considera que las dos esferas de acción anteriormente mencionadas ayudan a aumentar la resiliencia, a través de un mayor empleo, un acceso más equitativo a los recursos, una mayor protección frente a las situaciones de dificultad económica, o el progreso hacia la gobernabilidad democrática. Sin embargo, existen dos áreas de acción adicionales no necesariamente cubiertas con lo anterior y que merecen especial atención: la recuperación rápida y efectiva de las crisis en los casos en que haya fallado la prevención, y el fortalecimiento de la capacidad ante los desastres naturales agravados por el cambio climático.

### Gobernabilidad

La Asamblea General de NU y el Consejo Económico y Social establecen y aprueban las políticas del PNUD.

El órgano de gobierno del PNUD es su Junta Ejecutiva. La Junta Ejecutiva está compuesta por 36 miembros elegidos de manera rotatoria que representan a 5 grupos regionales. Entre las funciones de la Junta Ejecutiva se encuentra la de supervisar las actividades del PNUD, el progreso en su plan estratégico, y comprobar que efectivamente el PNUD responde a las necesidades de los países beneficiarios de los programas.

---

<sup>115</sup>Resiliencia entendida como capacidad de gestionar riesgos. Estos son concebidos como aquellos factores de magnitud e intensidad capaces de provocar a la vez una interrupción del progreso y el desarrollo y costos directos e indirectos importantes (DP/2013/40, Pág. 12)

El máximo responsable del PNUD es su administrador/a. Su puesto es designado directamente por el/la Secretario/a General por un periodo de 4 años, y confirmado por la Asamblea General<sup>116</sup>.

### Recursos

El PNUD recibe cerca de 9500 Millones US\$ por bienio de diferentes fuentes<sup>117</sup>:

- ✓ Contribuciones regulares de los Estados Miembros (20%)
- ✓ Contribuciones voluntarias para fines (proyectos) concretos procedentes de donantes bilaterales o multilaterales (65%)
- ✓ Recursos provenientes de los países beneficiarios de los programas del PNUD (*Cost sharing*) (15%)

### Función de evaluación

La función de evaluación en el PNUD se practica mayoritariamente de manera descentralizada, es decir las oficinas de campo realizan sus propios planes de evaluación, y contratan directamente personal externo como parte del equipo evaluador. La Oficina de Evaluación del PNUD, ubicada en la sede central del PNUD en Nueva York, es la responsable de velar por el funcionamiento de la evaluación en la organización.

La Oficina de Evaluación es una oficina independiente, que reporta a la Junta de Ejecutiva a través del Administrador/a del PNUD. Entre las responsabilidades de la Oficina de Evaluación se encuentran las siguientes: La realización de evaluaciones de carácter estratégico, la promoción del uso de la evaluación y lecciones aprendidas en la organización, y el desarrollo de procedimientos y guías metodológicas. La Oficina de Evaluación también participa activamente en diversas redes profesionales de evaluación y tiene un papel muy representativo dentro de UNEG al ostentar la secretaría.

El PNUD desarrolló su política de evaluación tras la publicación de las Normas y Estándares UNEG en 2005. La política de evaluación del PNUD fue aprobada por la Junta Ejecutiva en 2006.

### Proceso de revisión por pares

La Oficina de Evaluación del PNUD fue la primera organización voluntaria en someterse a un ejercicio de revisión por pares. La preparación comenzó a finales del 2004 y fue de aproximadamente un año. El panel de evaluación estaba compuesto por miembros de departamentos de cooperación bilateral danesa, británica y holandesa, FIDA y la red IDEAS. La revisión de la función de evaluación en el PNUD, llevada a

---

<sup>116</sup>[http://www.undp.org/content/undp/en/home/operations/about\\_us.html](http://www.undp.org/content/undp/en/home/operations/about_us.html)

<sup>117</sup> Fuente: Presupuesto Integrado PNUD 2014-2017, DP 2013/41

cabo en base a los criterios claves de utilidad, independencia y credibilidad, se realizó utilizando los siguientes métodos analíticos:

- ✓ Análisis documental extensivo de la información suministrada por la Oficina de Evaluación, y existente en la web o en diferentes repositorios.
- ✓ Muestra de evaluaciones gestionadas desde la Oficina de Evaluación. Estas incluyeron: La evaluación del segundo marco de cooperación global de PNUD y del Documento Estratégico de Reducción de la Pobreza entre otros.
- ✓ Entrevistas que incluyeron miembros de la Junta Ejecutiva, personal de la Oficina de Evaluación, y personal del PNUD.
- ✓ Reuniones con grupos focales para presentar y comentar los resultados preliminares en Noviembre de 2005.

El informe es públicamente accesible a través de la página web de UNEG (UNEG/UNDP 2006):

<http://www.uneval.org/document/detail/99>

Posteriormente en 2013, la Oficina de Evaluación del PNUD se sometió a una segunda revisión por pares con una metodología similar a la revisión anterior, aunque en este último caso la revisión estaba centrada en el análisis de la capacidad de la función de evaluación del PNUD para transferir conocimiento. En concreto, la credibilidad y la utilidad de la función de evaluación fueron los dos pilares analizados en esta segunda revisión por pares.

El informe es accesible a través de la página web de UNEG (UNEG/UNDP 2013)

<http://www.uneval.org/document/detail/1632>

Un año más tarde, en 2014, la función de evaluación del PNUD se somete de nuevo a un proceso de revisión de su política y prácticas evaluativas por parte de un grupo de profesionales independientes. En este caso, no se trata de un proceso de revisión por pares, puesto que no es un análisis realizado por otras agencias de NU bajo el marco de la UNEG. La metodología aplicada es muy similar a los procesos de revisión por pares, e incluye revisión documental, análisis de una muestra de informes de evaluación, entrevistas semi-estructuradas, encuestas electrónicas a los/as Representantes Residentes de PNUD, personal y consultores/as externos encargados de la realización de evaluaciones descentralizadas, y casos de estudio.

El informe es accesible a través de la página web del PNUD (Baastel 2014)

<http://web.undp.org/evaluation/>

## 1.3 PNUMA

### Mandato

El PNUMA es el programa de Naciones Unidas que tiene por responsabilidad principal la protección del medio ambiente. Su mandato proviene de la resolución de la Asamblea General de NU 2297 de Diciembre de 1972, a través de la cual se establece el Consejo Gobernador de PNUMA, su secretaría ejecutiva y el Fondo para el Medioambiente.

La misión de PNUMA es la de liderar y promover la búsqueda de alianzas en la protección del medioambiente a través de la información, diseminación y apoyo en la mejora de la calidad de vida de las personas sin comprometer las generaciones futuras.

Una función primordial de PNUMA es revisar las tendencias globales ambientales, detectar riesgos emergentes, y promover la cooperación internacional de manera coherente en materia medioambiental. Su mandato incluye el desarrollo de la agenda medioambiental a nivel global, y la promoción de una implementación coherente de la dimensión ambiental de desarrollo sostenible en el marco de trabajo de NU. Esto se resume en tres líneas principales de trabajo<sup>118</sup>:

- ✓ Analizar el estado y tendencia del medioambiente a nivel global, regional y nacional.
- ✓ Apoyar el desarrollo de política ambiental a nivel internacional y nacional.
- ✓ Fortalecer instituciones en sus responsabilidades de gestión ambiental.

### Marco Estratégico

La Estrategia de Medio Término (MTS) es el plan cuatrienal donde se establecen las líneas programáticas de acción de PNUMA y proporciona la visión y dirección del programa durante el periodo cubierto.

La Estrategia de Medio Término de 2014-2017<sup>119</sup> incluye seis prioridades temáticas:

- ✓ Cambio climático: Esta área se ha convertido en prioritaria para el conjunto de NU y autoridades gubernamentales. Sus efectos trascienden más allá de lo medioambiental, y requiere una intervención política, social y económica. El cambio climático supone un riesgo para los avances realizados en desarrollo humano a nivel global. El PNUMA contribuye a los esfuerzos contra el cambio climático a través de tres pilares: Adaptación, Mitigación y Reducción de emisiones por deforestación y degradación de bosques, los dos primeros, definidos en la Convención Marco de Naciones Unidas contra el Cambio Climático (CMNUCC). En este sentido, PNUMA colabora con los países para reducir su huella de carbono e incrementar sus niveles de

---

<sup>118</sup>[http://www.pnuma.org/PNUMA\\_ORPALC.pdf](http://www.pnuma.org/PNUMA_ORPALC.pdf)

<sup>119</sup>UNEP/GC.27/9

resiliencia ante los impactos del cambio climático, a través de la sensibilización, intercambio de información, evaluación de necesidades tecnológicas, e implementación de proyectos para la mitigación y adaptación.

- ✓ Desastres y conflictos. La explotación sin control de recursos naturales, y las consecuencias del cambio climático en cuanto a la disponibilidad de agua y seguridad alimentaria, pueden agravar conflictos sociales y políticos. Hasta el momento, las guerras civiles se han centrado en recursos de alto valor como minerales y petróleo; otros conflictos pueden tener su origen en el control de los recursos como agua y tierras fértiles. Los factores ambientales son rara vez la única causa de conflictos violentos. PNUMA colabora con los países en la identificación y puesta en marcha de medios de subsistencia alternativos y sostenibles con el medioambiente que prevengan conflictos relacionados con la gestión de los recursos.
- ✓ Gestión del ecosistema. Los avances tecnológicos nos permiten hacer una mayor optimización de la gestión de nuestros recursos. En este sentido, es fundamental que comunidades locales, empresas y entidades gubernamentales ayuden de manera coherente a innovar en beneficio de las comunidades, sus economías y el medio ambiente global. El PNUMA colabora con las instituciones en la articulación de medidas coherentes para el uso y la gestión de los recursos. Esto incluye, entre otras medidas, la optimización de prácticas de riego, determinación de balances hídricos en cuencas, y la gestión integrada de zonas costeras.
- ✓ Gobernabilidad ambiental. Esta área prioritaria se enfoca en apoyar el desarrollo de políticas y el fortalecimiento del marco regulatorio a nivel nacional, regional y mundial para atender las prioridades ambientales. Entre las principales líneas de acción, se contemplan las siguientes: 1) la promoción del desarrollo progresivo y la implementación del derecho ambiental para dar respuesta efectiva a los problemas ambientales y 2) El apoyo a la implementación de los acuerdos multilaterales ambientales.
- ✓ Productos químicos y desechos. Es evidente que la gestión inadecuada de productos químicos y desechos puede convertirse en una seria amenaza para la salud de las personas y el medio ambiente. Para prevenir estos riesgos, los países han realizado esfuerzos en el desarrollo de políticas y regulaciones, pero todavía queda mucho por hacer. PNUMA apoya las iniciativas gubernamentales orientadas a la mejora en la gestión de los productos químicos y desechos, y proporciona asistencia técnica para la implementación de los acuerdos internacionales vigentes o en fase de negociación.
- ✓ Eficiencia y sostenibilidad en la producción y el consumo. Dentro de esta área prioritaria, se pretende asegurar que los recursos naturales sean procesados y consumidos de una forma más ambiental y sostenible. El PNUMA está

colaborando con los países a través de asistencia técnica en la búsqueda de una mayor eficiencia en la gestión de los recursos, consumo sostenible y patrones de producción.

Cada una de las líneas prioritarias establecidas en la MTS esta desglosada en objetivos y resultados esperados. Los Programas de Trabajo (POW) son aprobados de manera bianual con el fin de operativizar las líneas estratégicas en programas y proyectos. Actualmente está en vigor el POW 2014-2015<sup>120</sup>.

### Gobernabilidad

La estructura de órganos de gobierno de PNUMA es la siguiente<sup>121</sup>:

- ✓ El Consejo de Gobierno (GC) es el encargado de establecer y aprobar las líneas estratégicas de PNUMA, incluyendo presupuesto, Estrategia de Medio Término (MTS) y Programa de Trabajo (POW). Está compuesto por 58 miembros elegidos por la Asamblea General de NU que representan a diferentes grupos regionales en periodos rotatorios de cuatro años.
- ✓ El Comité de Representantes Permanentes (CPR), que actúa como órgano de control y supervisión para el cumplimiento de las directrices establecidas por el Consejo de Gobierno. El CPR funciona a través de subcomisiones o grupos de trabajo, y la frecuencia de sus reuniones se establece en función de sus necesidades y plan de trabajo.
- ✓ La Secretaría del PNUMA, que se encarga de ejecutar las decisiones del Consejo de Gobierno y hacerlas operativas a través de estrategias, programas y proyectos. Su sede está en Kenia. La Secretaría se compone de: Oficinas, encargadas de áreas de gestión y apoyo; Divisiones Temáticas, responsables de suministrar conocimiento técnico sobre un tema, Oficinas regionales y Oficinas de Enlace, que ejecutan el programa de trabajo a nivel regional o de país. El/la Directora/a Ejecutivo/a es el/la máximo/a responsable de la Secretaría.

### Recursos

El presupuesto bianual del PNUMA fue de 415 Millones de US\$ en el bienio 2010-2011<sup>122</sup>. Las fuentes de financiación de PNUMA provienen de tres fuentes principales:

- ✓ Contribuciones regulares de los Estados Miembros (14 Millones)

---

<sup>120</sup> UNEP/GC 27.10 Proposed biennial programme of work and budget for 2014-2015.

<sup>121</sup><http://www.unep.org/about/sgb/Default.aspx>

<sup>122</sup>Fuente: MTS UNEP GC.27/9

- ✓ Contribuciones extrapresupuestarias (228 Millones). Aquí se incluyen las contribuciones de los gobiernos, organizaciones privadas, etc...a actividades específicas del programa de trabajo, generalmente a través de fondos fiduciarios. Los fondos fiduciarios se dividen entre los fondos fiduciarios generales, y los de cooperación técnica. De entre los generales, el relativo a la Convención de Diversidad Biológica y el de la Protección del Mar Mediterráneo son los más relevantes. Entre los fondos fiduciarios de cooperación técnica, el del FMAM (GEF) es el más importante. PNUMA es una agencia implementadora de fondos FMAM, y por lo tanto puede solicitar y recibir dichos fondos para el diseño e implementación de proyectos.
  
- ✓ El Fondo de Medioambiente (180 Millones). Este Fondo, establecido en 1972 es la principal fuente de contribuciones voluntarias y no asignadas a proyectos concretos (non-earmarked funding). Con ello, se cubren los gastos de las unidades administrativas de la secretaría, y actividades prioritarias de lo que constituye el programa de trabajo.

### Función de Evaluación

La Oficina de Evaluación es la unidad dentro de la secretaría con el mandato de implementar la función de evaluación en PNUMA. Su ámbito de actuación cubre todos los programas y proyectos financiados por el Fondo de Medioambiente, y otros fondos fiduciarios, incluidos el fondo de cooperación técnica del FMAM. La Oficina de Evaluación pertenece a la Oficina Ejecutiva, reporta al Director Ejecutivo, y a través del Director Ejecutivo, a los órganos de gobierno (El Consejo Ejecutivo y el Consejo de Representantes Permanentes).

Entre el tipo de actividades que realiza la Oficina de Evaluación destacan las evaluaciones de proyecto, las evaluaciones de los subprogramas de los Programas de Trabajo (POW), el control de la calidad en la formulación y gestión de los proyectos, evaluaciones de áreas específicas de la Estrategia de Medio Término, estudios relacionados con la gestión (análisis de herramientas internas de gestión en la organización), y meta-evaluaciones.

En promedio la Oficina de Evaluación realiza entre 15 y 20 evaluaciones anuales, en su mayoría cuando los proyectos terminan, aunque en ocasiones también se realizan evaluaciones de medio término cuando se observa dificultades durante la implementación de los proyectos, y en proyectos implementados junto con otras agencias de NU<sup>123</sup>.

---

<sup>123</sup> UNEG/UNEP2012 y <http://www.unep.org/eou/>

### Proceso de revisión por pares

El proceso de revisión por pares se llevó a cabo entre Septiembre de 2011 y Marzo de 2012. El panel, compuesto por evaluadores profesionales de otras agencias de NNUU y organismos bilaterales de cooperación realizó una extensa revisión documental de la organización y estructura de PNUMA y de su función de evaluación, y entrevistas a miembros de los órganos de gobierno, equipo de la Oficina de Evaluación, y managers de las unidades técnicas de PNUMA. Esta información fue complementada con entrevistas telefónicas efectuadas a los consultores que han colaborado en los equipos de evaluación de PNUMA, y otras agencias externas, como el Director de la Oficina de Evaluación del FMAM.

El informe es accesible a través de la página web del UNEG (UNEG/UNEP2012)

<http://www.uneval.org/document/detail/1163>

## 1.4 ONUDI

### Mandato

La ONUDI fue establecida en 1966 a través de la resolución de la Asamblea General 2152 (XXI) de 17 de Noviembre del mismo año. Sin embargo no fue hasta 1985 cuando ONUDI se convierte en agencia especializada, por la resolución 3362 (S-VII) de la Asamblea General. La Carta de Constitución de UNIDO entró en vigor en Junio de 1985, y la primera sesión de la Conferencia General tuvo lugar en Agosto de 1985.

La misión de la ONUDI dentro del marco de trabajo de NU es la de promover y acelerar la industrialización en los países en desarrollo. La ONUDI aspira a la reducción de la pobreza a través del desarrollo industrial sostenible. Para ello, proporciona apoyo y asistencia técnica para el desarrollo del sector productivo, la participación en el comercio internacional y la protección del medio ambiente.

En 2013 tuvo lugar la Declaración de Lima<sup>124</sup>, donde la Conferencia General de la ONUDI destacó la importancia del desarrollo industrial 'inclusivo y sostenible'. El desarrollo industrial inclusivo y sostenible implica:

- ✓ Aumentar el nivel de industrialización de las economías y obtener un mayor beneficio de la globalización de los mercados de bienes y servicios industriales.
- ✓ El crecimiento industrial debe ser inclusivo y compartido igualmente entre hombres y mujeres.
- ✓ El crecimiento industrial debe ser medioambientalmente sostenible.
- ✓ El conocimiento y la optimización de los recursos ayudan a maximizar el impacto del desarrollo industrial sostenible e inclusivo.

A través de esta declaración, se reconoce que la erradicación de la pobreza sigue siendo una necesidad central, que solo es superable mediante un crecimiento económico e industrial fuerte, inclusivo y sostenible. Para ello, se requiere una integración eficaz de las dimensiones económica, social y medioambiental del desarrollo sostenible.

Esta declaración resulta un importante apoyo a la organización, la cual ha sufrido relevantes críticas e importantes abandonos, que incluyen a Australia, Canadá y EEUU, y más recientemente Reino Unido, Francia y Bélgica.

### Marco Estratégico

El marco de trabajo se establece bianualmente en el documento Programa y Presupuesto y es aprobado por la Conferencia General. De acuerdo al Programa y Presupuesto en vigor, las tres áreas prioritarias de trabajo de ONUDI incluyen<sup>125</sup>:

---

<sup>124</sup>[http://www.unido.org/fileadmin/user\\_media\\_upgrade/Who\\_we\\_are/Structure/Polymaking\\_Organs/Lima\\_Declarati on\\_EN\\_web.pdf](http://www.unido.org/fileadmin/user_media_upgrade/Who_we_are/Structure/Polymaking_Organs/Lima_Declarati on_EN_web.pdf)

<sup>125</sup><http://www.unido.org/en/what-we-do.html>

- ✓ Reducción de la pobreza mediante actividades productivas. A través de esta prioridad temática, la ONUDI contribuye a actividades orientadas a alcanzar las metas propuestas en los Objetivos de Desarrollo del Milenio (ODM), puesto que la actividad productiva tiene un papel fundamental en la reducción de la pobreza. En este sentido, se promueve la cooperación entre los sectores público y privado. Las autoridades gubernamentales tienen la responsabilidad de regulación y formulación de políticas, y el sector privado ponen en marcha sus actividades a través de pequeñas y medianas empresas (PYME).
- ✓ Creación de capacidad comercial. Esta área prioritaria tiene como propósito fortalecer la capacidad de los países en desarrollo y economías en transición para participar en el comercio mundial como parte integrante del crecimiento económico de estos países. ONUDI presta apoyo a las empresas para ofrecer productos competitivos, fiables y costo-efectivos. Para ello, es necesario previamente identificar sectores y productos con potencial competitivo y formular políticas destinadas a aumentar la competitividad industrial, superar los obstáculos técnicos al comercio y cumplir con las medidas sanitarias y fitosanitarias. Además, ONUDI apoya los procesos de modernización de los procesos de fabricación en sectores con potencial de exportación y apoya la creación de consorcios de exportación – un tipo especializado de red de PYMES – y el diseño e implementación nacional y regional de programas de creación de capacidad comercial centrada en productos básicos. Estas actividades se llevan a cabo en cooperación con organismos asociados internacionales como el Centro Internacional de Comercio (UNCTAD/OMC), la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), el Fondo Común para los Productos Básicos (FCPB), y la Organización Mundial del Comercio (OMC).
- ✓ Medioambiente y energía. ONUDI apoya la promoción de métodos de producción más limpia en todo el sistema de desarrollo industrial. La promoción de la eficiencia en el uso de recursos exige tomar en consideración tanto el valor económico como la sostenibilidad del medio ambiente. Entre los servicios de asistencia técnica que la ONUDI proporciona, se incluyen: el uso racional de los recursos y la producción industrial con bajas emisiones carbónicas; la energía renovable para usos productivos; y el apoyo a la aplicación de los acuerdos multilaterales sobre el medio ambiente.

## Gobernabilidad

UNIDO tiene dos órganos de gobierno principales: La Conferencia General y la Junta de Desarrollo Industrial. El Comité de Programa y Presupuesto es un órgano subsidiario de la Junta de Desarrollo Industrial<sup>126</sup>.

---

<sup>126</sup><http://www.unido.org/en/who-we-are/structure.html>

- ✓ La Conferencia General (GC), formada por 169 Estados Miembros<sup>127</sup>, determina los principios y políticas estratégicas de la organización y aprueba los presupuestos y planes de trabajo. La Conferencia General elige cada cuatro años al Director/a General. También elige a los representantes de la Junta de Desarrollo Industrial y del Comité de Programa y Presupuesto. Se reúne de manera bianual.
- ✓ La Junta de Desarrollo Industrial (IDB) se compone de 53 miembros, elegidos de manera rotatoria cada 4 años. Revisa la implementación del plan de trabajo, el presupuesto, y prepara recomendaciones para la atención de la Conferencia General, incluyendo la elección del Director General de la ONUDI. Se reúne anualmente.
- ✓ El Comité de Programa y Presupuesto lo componen un total de 27 miembros, elegidos de manera bianual. Es un órgano subsidiario y operativo de la Junta de Desarrollo Industrial, y le proporciona apoyo en la preparación del plan de trabajo, presupuesto y otros temas organizacionales. Se reúne anualmente.

El/la directora/a General es la responsable de la administración y gestión de la ONUDI ante los órganos de gobierno. Es elegido/a por la Conferencia General en base a la recomendación de la Junta de Desarrollo Industrial por un periodo de cuatro años.

Existen 6 unidades independientes que reportan directamente al Director General: La Oficina de Planificación Estratégica, Coordinación, y Monitoreo de Calidad, la Oficina del Asesor Especial, Oficina de Asuntos Legales, la Oficina de la Evaluación Independiente, La Oficina de Control y Supervisión Interna, y los Puntos Focales de Ética y Rendición de Cuentas.

Además de la sede en Viena, UNIDO tiene oficinas en Bruselas, Ginebra y Nueva York. La red de ONUDI en el campo incluye 30 oficinas regionales y nacionales y 17 puntos focales que cubren un total de 116 países. ONUDI también tiene oficinas para la Promoción de Inversiones y Tecnología en 6 países. A su vez, apoya un total de 56 Centros y Programas de Producción Más Limpia en cooperación con el Programa de Naciones Unidas para el Medioambiente.

## Recursos

El Programa y Presupuesto para el bienio 2013-2014 suma un total de 500 Millones de US\$, de los cuales 350 Millones están destinados a proyectos de cooperación técnica. A continuación se resume la procedencia de los fondos para los proyectos de cooperación técnica en millones de US\$:

---

<sup>127</sup>A fecha de Mayo de 2015.

**Tabla A.1.1: Recursos presupuestados a cooperación técnica 2013-2014**

	<b>2013</b>	<b>2014</b>
Presupuesto ordinario	6,7	6,2
PNUD	0,4	0,8
Fondo para el Desarrollo Industrial	22,6	22,2
Protocolo de Montreal	58,9	38,8
Fondo fiduciario autofinanciado	6,2	3,6
Fondos fiduciarios financiados por terceros	32,6	28,8
Fondo para el Medioambiente	32,5	45,7
Otros Fondos fiduciarios	9,1	4,9
UE	11,5	20,2

Fuente: Informe Anual de ONUDI 2014

### Función de Evaluación

La Unidad de Evaluación (EVA) es la responsable por la función de evaluación en la ONUDI. La Política de Evaluación de la ONUDI, aprobada en 2006, es el marco que rige las actividades de evaluaciones. Entre las responsabilidades de EVA, se incluyen las siguientes:

- ✓ Realizar evaluaciones independientes y dar seguimiento a las recomendaciones de los de informes de evaluación. Estas evaluaciones pueden ser a nivel de programas, proyectos, temáticas, e incluso cubrir procesos de gestión interna.
- ✓ Promover el uso de los hallazgos, lecciones aprendidas y recomendaciones, en la programación y formulación de nuevos programas.
- ✓ Desarrollar guías, metodología y procedimientos para mejorar la calidad de los proyectos, y favorecer su alineamiento con un enfoque orientado a resultados.
- ✓ Preparar informes sobre actividades de evaluación para los órganos de gobierno.
- ✓ Desarrollar y promover alianzas con otras organizaciones del sistema de NU en el ámbito UNEG y OCDE-DAC para intercambiar experiencias relacionadas con evaluación y favorecer la disseminación de buenas prácticas

La unidad de evaluación EVA está ubicada dentro de la Oficina para el Aprendizaje y la Estrategia Organizacional, y, por tanto, es independiente de las funciones temáticas y gerenciales responsables de la implementación y gestión de proyectos. El Director/a de EVA es nominado/a por el/la director/a general.

## Proceso de Revisión por pares

El proceso de revisión por pares tuvo una duración total de 1 año, desde Abril de 2009 a Abril de 2010. La metodología empleada incluye una amplia revisión documental sobre procedimientos y estructura organizacional en la ONUDI, estudios de casos a cuatro evaluaciones realizadas por EVA<sup>128</sup>, una encuesta a managers y responsables de las unidades técnicas, entrevistas a personal responsable de tareas administrativas y operacionales, y dos discusiones de grupos focales.

Uno de los grupos focales estuvo compuesto por directores/as operacionales que se enfocaron en el rol de EVA, procesos de evaluabilidad y gestión del conocimiento. El otro grupo focal estuvo formado por todos directores/as y personal entrevistado a lo largo del proceso de revisión por pares, donde se presentaron y discutieron los hallazgos y principales conclusiones del proceso de revisión por pares.

El informe es públicamente accesible a través de UNEG (UNEG/UNIDO 2010):

<http://www.uneval.org/document/detail/655>

---

<sup>128</sup>India Country Service Framework, Burkina Faso Integrated Programme, Investment and Technology Promotion Office and UNIDO-UNEP Cleaner Production Programme.

## 1.5 UNICEF

### Mandato

UNICEF se crea en 1946 para prestar asistencia a los millones de niños/as que sufren las graves consecuencias de la Segunda Guerra Mundial. En 1953 adquiere el estado de organismo permanente de las NU y se amplía su mandato a una visión más duradera y multifacética de apoyo a la niñez, tanto a nivel educacional, como social y sanitario. Cinco años más tarde, en 1959 se aprueba la Declaración de los Derechos del Niño, donde se ratifica la atención y el enfoque a los derechos de los niños/as en educación, atención a la salud y nutrición adecuada.

La labor de UNICEF se expande geográficamente y en la década de los 60 UNICEF es un claro referente mundial. En 1965, obtiene el Premio Nobel de la Paz por su labor en pro de la confraternidad entre las naciones. La Convención sobre los Derechos del Niño es aprobada por la Asamblea General de NNUU en 1989. Un año más tarde se celebra la Cumbre Mundial en favor de la Infancia en Nueva York en 1990 donde se establecen metas para el decenio en materia de educación, salud y nutrición de los niños/as, y donde UNICEF es referente para su coordinación y seguimiento, como lo es también de algunos de los indicadores de los Objetivos de Desarrollo del Milenio, enunciados en la Declaración del Milenio, firmada en 2000.

La misión de UNICEF consiste en promover los derechos de los niños/as desde el concepto de equidad. La estrategia de equidad es entendida como *“la oportunidad de que todos los niños/as tengan la oportunidad de sobrevivir, desarrollar y alcanzar su máximo potencial sin discriminaciones ni prejuicios”* (UNICEF Plan estratégico 2014-2017, pág. 2)

### Marco estratégico

El Plan Estratégico es el documento cuatrienal que dirige las prioridades y líneas estratégicas de UNICEF. Actualmente está en vigor el Plan Estratégico 2014-2017 donde se señalan siete áreas de acción<sup>129</sup>:

- ✓ Salud: UNICEF, en colaboración con la Organización Mundial de la Salud (OMS) y otras entidades, presta apoyo a los países para la puesta en marcha de políticas y programas orientados a reducir la mortalidad infantil. Esto incluye la participación en el Plan Mundial de Erradicación de la Poliomielitis 2013-2018, y atención en el tratamiento de intervenciones preventivas, fortalecimiento de los sistemas de salud, y la mejora de la calidad en la gestión y utilización de datos en los servicios de salud, con principal atención a las madres, los/as recién nacidos/as y los/as niños/as.
- ✓ VIH y SIDA: Apoyo y asistencia técnica en las intervenciones de prevención y tratamiento del VIH en niños/as, mujeres embarazadas y adolescentes. UNICEF

---

<sup>129</sup>E/ICEF/2013/21

coordina el Plan Mundial para eliminar las nuevas infecciones por VIH. También apoya actividades de promoción de una educación sexual integral y en proteger los derechos de los/as niños/as, mujeres embarazadas y adolescentes en riesgo de discriminación por estar afectados/as por el VIH.

- ✓ Agua, saneamiento e higiene: A través de esta área prioritaria se promueve la utilización equitativa del agua potable, saneamiento y mejora en las prácticas de higiene. Esto incluye apoyo a los gobiernos para mejorar el acceso sostenible al agua potable, el acceso a servicios de saneamiento adecuados, diseminación de buenas prácticas de higiene e instalación de servicios básicos sanitarios en escuelas y centros de salud. UNICEF implementa en coordinación con el Banco Mundial, entre otras instituciones, la estrategia 'saneamiento y agua para todos'<sup>130</sup>.
- ✓ Nutrición: Apoyo y asistencia a los gobiernos en la implementación de políticas y estrategias de mejora nutricional. Se incluye aquí el suministro de suplemento de vitaminas y micronutrientes; promoción de la lactancia materna; y la prevención y gestión comunitaria de la desnutrición infantil. También se incluye en esta área prioritaria el apoyo a las instituciones gubernamentales en la protección de la situación nutricional de los/as niños/as en situaciones de crisis humanitarias.
- ✓ Educación: El Plan Estratégico incluye la educación equitativa e inclusiva como área prioritaria. En este sentido, se apoya el fortalecimiento de los sistemas educativos que promueven la integración de los/as niños/as con discapacidad, enfatizar la importancia de la educación en la consolidación de la paz, el monitoreo de la matriculación escolar en la edad correcta, y en general, reforzar la capacidad de los gobiernos e instituciones a nivel local para mejorar el acceso y la calidad de la educación.
- ✓ Protección a la niñez: Esta área desarrolla el apoyo a las estrategias de prevención y respuesta ante situaciones de malos tratos, explotación y abandono infantil. Esto incluye una mayor atención a la interacción entre los sistemas de protección de la infancia y las normas sociales, y el fortalecimiento de la capacidad de protección de las familias desde un enfoque intersectorial. También incluye las medidas de preparación y respuesta ante situaciones de emergencia, y conflictos armados.
- ✓ Inclusión social: Apoyo a los países en el diseño e implementación de sistemas de protección social, en colaboración con la Organización Internacional del Trabajo (OIT). Esto incluye la formulación de políticas que promuevan la inclusión, y su alineación con los sistemas presupuestarios en el área social. Para ello, UNICEF colabora con los gobiernos en su condición de garantes de derechos, y con las familias en el cuidado de los/as niños/as para que puedan conocer, solicitar y acceder a los servicios básicos de protección social.

---

<sup>130</sup> <http://sanitationandwaterforall.org/about/>

Estas siete áreas prioritarias estén intrínsecamente relacionadas, y, por tanto, su implementación requiere de programas y proyectos que integren y faciliten las sinergias de los diferentes resultados.

### Gobernabilidad

El órgano de gobierno de UNICEF es su Junta Ejecutiva<sup>131</sup>. Está integrada por 36 miembros representantes de los 5 grupos regionales y elegidos por el Consejo Económico y Social de las NNUU por periodos de 3 años. La Junta Ejecutiva es la responsable de orientar las actividades de UNICEF, examinar y aprobar su Plan Estratégico y el presupuesto. La Junta Ejecutiva se reúne 3 veces al año, y está apoyada por la Mesa y la Secretaría de la Junta Ejecutiva.

La Mesa de la Junta Ejecutiva coordina la labor de la Junta Ejecutiva y actúa de enlace entre la Secretaría de UNICEF y los grupos regionales representados en la Junta Ejecutiva. Facilita la negociación de las decisiones en el periodo de sesiones. La Mesa se reúne mensualmente y en los periodos de sesiones de la Junta Ejecutiva. Está integrada por el Presidente y cuatro Vicepresidentes, representantes de los grupos regionales. La Secretaría de la Junta Ejecutiva es responsable de las tareas administrativas que cubren las actividades de enlace entre la Junta Ejecutiva y la Secretaría de UNICEF. Incluye la preparación y coordinación de reuniones y consultas, y visitas sobre el terreno de los miembros de la Mesa y de la Junta Ejecutiva, y servicios documentales.

UNICEF es una organización con fuerte presencia en el terreno y estructuras descentralizadas. La Secretaría de UNICEF está en Nueva York, desde donde se formulan las políticas temáticas, y se gestiona y coordinan las actividades. Cuenta con siete oficinas regionales, y oficinas en el terreno en más de 190 países. La máxima responsabilidad de la Secretaría de UNICEF la ostenta el/a Director/a Ejecutivo/a.

UNICEF también gestiona el Centro de Investigaciones Innocenti en Florencia (Italia) y las Oficinas en Tokio (Japón), y Bélgica (Bruselas), donde se realizan actividades de recaudación de fondos, y sirven de enlace con las divisiones centrales de formulación de políticas. A su vez, la labor de UNICEF se apoya en los Comités Nacionales, que son organizaciones no gubernamentales que promueven las actividades de UNICEF, recaudan fondos y establecen contactos y alianzas con el sector privado y la sociedad civil. Se estima que estos Comités de UNICEF recaudan un tercio de los recursos de UNICEF<sup>132</sup>.

---

<sup>131</sup><http://www.unicef.org/about/execboard/>

<sup>132</sup>[http://www.unicef.org/about/structure/index\\_natcoms.html](http://www.unicef.org/about/structure/index_natcoms.html)

## Recursos

UNICEF se financia exclusivamente de contribuciones voluntarias, de las cuales cerca de un tercio son recaudadas a través de los 36 Comités Nacionales. Las contribuciones provienen de empresas privadas, organizaciones locales y donantes.

El presupuesto aprobado para el periodo cubierto por el Plan Estratégico 2014-2017 se eleva a 2.094,5 millones de US\$. El presupuesto institucional es 97,0 millones de dólares, un 4,8%, superior al aprobado en el período 2010-2013. El presupuesto institucional comprende los siguientes recursos: 567,5 millones de dólares para la eficacia de las actividades de desarrollo; 36,6 millones para las actividades de coordinación de las Naciones Unidas en favor del desarrollo; 1.435,4 millones para la gestión; y, dentro de los fines especiales, 55,0 millones para inversión en capital<sup>133</sup>.

## Función de evaluación

La función de evaluación en UNICEF se caracteriza por estar principalmente descentralizada a las oficinas en el terreno. Se estima que cerca de 100 evaluaciones se realizan al año, bien sean de programas, proyectos, temáticas o de procesos.

La primera Política de Evaluación de UNICEF fue aprobada por la Junta Ejecutiva en 2008 (E/ICEF/2008/4). Posteriormente, ha sido actualizada en 2013 (E/ICEF/2013/14). La Política de evaluación delinea los propósitos de la evaluación en UNICEF, y los roles de coordinación de la unidad de evaluación central con las oficinas nacionales y regionales.

La unidad de evaluación central establece la metodología, los procedimientos, y los controles de calidad para garantizar el funcionamiento de la evaluación en el terreno. También realiza evaluaciones temáticas, cuyo plan de trabajo está vinculado al Plan Estratégico de UNICEF. El/la directora/a de Evaluación presenta anualmente a la Junta Ejecutiva el informe anual de evaluación, que incluye análisis de síntesis y calidad de las evaluaciones. Por su parte, las oficinas nacionales realizan sus planes de trabajo y estos son aprobados a nivel nacional y regional. UNICEF tiene un papel representativo y alto liderazgo en la promoción de la evaluación y buenas prácticas en UNEG, a través de su participación en los diferentes grupos de trabajo<sup>134</sup>.

## Proceso de revisión por pares

UNICEF fue la segunda agencia, después de PNUD, en someterse al proceso de revisión por pares. El proceso tuvo una duración total de un año, que culminó con la presentación de los principales hallazgos y recomendaciones a la Junta Ejecutiva de UNICEF en Junio de 2006.

---

<sup>133</sup>Fuente: Presupuesto Integrado de UNICEF para 2014-2017, E/ICEF/2013/AB/L.4

<sup>134</sup>UNICEF Evaluation Brief 2014

La función de evaluación se revisó en torno a tres criterios: Independencia, credibilidad y utilidad. El panel estuvo compuesto por 8 evaluadores/as profesionales de organismos de NU y cooperación para el desarrollo bilateral (Canadá, Irlanda, Noruega) y multilateral (Banco Africano de Desarrollo, y la Asociación de Evaluación de Desarrollo Internacional).

La metodología incluye amplia revisión documental, entrevistas a personal de UNICEF a nivel central y en las oficinas nacionales, miembros de la Junta Ejecutiva, personal de la unidad de evaluación y consultores/as externos/as. También se realizó un estudio de caso, y, para ello, se seleccionó un país (Ghana) donde se analizó en detalle los procesos y el sistema existente que guía la función de evaluación descentralizada en UNICEF.

El informe ejecutivo es públicamente accesible en la web de UNEG (UNEG/UNICEF 2006):

<http://www.uneval.org/document/detail/98>

El panel de la revisión por pares hizo una serie de recomendaciones para la puesta en marcha de medidas de fortalecimiento de la función de evaluación: Entre ellas, destaca la de la adopción de una Política de Evaluación en línea con las normas y estándares UNEG, procedimientos específicos para una mayor predictibilidad del presupuesto para evaluación, y mejorar la capacidad y apoyo técnico a las oficinas de campo en monitoreo y evaluación.

Posteriormente, en Febrero de 2013, consultores/as independientes han realizado una revisión de la función de evaluación. El estudio no forma parte de lo que se considera los informes de revisión por pares, puesto que no se ha realizado por paneles de evaluación en el marco de UNEG, pero su contenido también es relevante para analizar la evolución de la función de evaluación a partir de las conclusiones de la revisión por pares.

El informe es públicamente accesible a través de la página web de UNICEF (Universalía 2013):

[http://www.unicef.org/evaluation/files/UNICEF\\_Synthesis\\_Report\\_Revised\\_Final.pdf](http://www.unicef.org/evaluation/files/UNICEF_Synthesis_Report_Revised_Final.pdf)

## 1.6 FAO

### Mandato

FAO fue establecida como organismo independiente de NNUU en 1945, con el mandato de apoyar a los países para que nunca sufran sus habitantes el problema del hambre, ni la malnutrición. Con el tiempo, el trabajo de la organización se abre a nuevas áreas de actuación, entre las que se encuentran los códigos y estándares sanitarios, la pobreza y el desarrollo rural, el cambio climático y la sostenibilidad de los recursos naturales. Por ejemplo, en 1962 se crea la Comisión del Codex Alimentarius FAO/OMS<sup>135</sup> para el establecimiento y regulación de normas alimentarias.

Años más tarde, en 1974, la Conferencia Mundial de la Alimentación, celebrada en Roma, recomienda la adopción de un compromiso internacional sobre seguridad alimentaria mundial. La relevancia de su mandato da un papel de referente a la FAO en agricultura. En 1992, la FAO y la OMS convocan la Conferencia Mundial sobre la Nutrición. En 1996, FAO organiza la Cumbre Mundial sobre la Alimentación. Como resultado, los jefes de Estado adoptan la Declaración de Roma sobre la Seguridad Alimentaria Mundial y el Plan de Acción de la Cumbre Mundial sobre la Alimentación. En el 2002, cinco años después de la Cumbre Mundial de la Alimentación, se confirma el compromiso internacional de reducir el número de personas subalimentadas a la mitad para el 2015, y en el 2004 el Consejo de la FAO adopta las directrices voluntarias en apoyo de la realización progresiva del derecho a una alimentación adecuada en el contexto de la seguridad alimentaria nacional.

El papel relevante de la FAO en seguridad alimentaria es incuestionable, pero también se ha debido adaptar a la multiplicación de otras organizaciones y sectores altamente comprometidos con la investigación y análisis de la agricultura a nivel global. Esto incluye por ejemplo a CGIAR (Consultative Group on International Agricultural Research), el sector privado y numerosas ONGs con mandatos vinculados a la seguridad alimentaria y agricultura.

En este sentido, FAO se ha sometido recientemente a un profundo proceso de reformas, fruto del Plan de Acción Inmediato aprobado por la Conferencia de la FAO en Noviembre de 2008<sup>136</sup>. Este Plan tiene como objetivo posicionar de una manera más estratégica el mandato de la FAO, y buscar alianzas sinérgicas con aquellos actores clave del sector privado, agencias bilaterales y ONGs que trabajan en favor de una agricultura sostenible.

---

<sup>135</sup> <http://www.codexalimentarius.org/codex-home/es/>

<sup>136</sup>FAO, C/2008/LIM/3 Plan de Acción Inmediato

## Marco Estratégico

En Enero de 2012, el Sr. José Graziano da Silva toma posesión como Director General de la FAO, y continua con el proceso de reflexión y reformas en la FAO con el objeto de determinar su orientación estratégica. Este proceso culmina con la revisión del marco estratégico para 2010-2019, que fue aprobado en Junio de 2013 por la Conferencia General<sup>137</sup>. El marco estratégico revisado propone cinco objetivos:

- ✓ Eliminar el hambre, la inseguridad alimentaria y la malnutrición. Esta área prioritaria necesita de una clara voluntad y compromiso político y una acción más coordinada para la mejora de la seguridad alimentaria y el estado nutricional de la población. FAO apoya a los países miembros en el establecimiento y la promoción de compromisos políticos explícitos, de mecanismos adecuados de gobernabilidad a nivel mundial, regional y nacional, y en el fortalecimiento de capacidades para la rendición de cuentas y supervisión de políticas, programas e inversiones sectoriales e intersectoriales.
- ✓ Aumentar la productividad de la agricultura, la actividad forestal y la pesca. Con el fin de satisfacer la demanda creciente, los sectores de la agricultura, actividad forestal y pesca debe realizar una transición hacia sistemas más sostenibles de producción y consumo. FAO apoya a los gobiernos en la implementación de prácticas más sostenibles, establecimiento de mecanismos de gobernanza viables, y propuestas de acción intersectorial. Esto supone apoyar la integración de los tres pilares básicos de sostenibilidad (ambiental, económico y social), reflejando el resultado y compromiso de la Conferencia de Rio + 20<sup>138</sup>.
- ✓ Reducir la pobreza rural. Los gobiernos deben seguir apoyando el desarrollo de políticas orientadas a favorecer el desarrollo rural y la reducción de la pobreza en el campo. El trabajo de la FAO se centra en: apoyar la orientación de políticas para una gestión conjunta de los recursos, y la integración de las organizaciones de productores; favorecer las oportunidades el empleo rural en condiciones dignas; y potenciar las sinergias y alineamiento entre las medidas de protección social y los programas de desarrollo rural.
- ✓ Fomentar sistemas agrícolas y alimentarios integradores y eficientes. Esta área prioritaria se enfoca en el fortalecimiento de capacidades a nivel local, la participación del sector privado en dichos sistemas y apoyar medidas para unos mercados justos y eficientes. Para ello, FAO apoya técnicamente el establecimiento de normas y servicios de regulación; proporciona recomendaciones para el funcionamiento eficiente de los mercados, y analiza los esfuerzos de inversión conjunta entre los sectores público y privado.

---

<sup>137</sup>FAO, C/2013/7

<sup>138</sup><http://www.uncsd2012.org/>

- ✓ Incrementar la resiliencia de los medios de vida ante las catástrofes. Esta área incluye la necesidad de reducir los riesgos, y a su vez, promover la preparación y los mecanismos de recuperación con el fin de facilitar las sinergias entre los esfuerzos humanitarios, de desarrollo e inversión. Con este objetivo, FAO presta apoyo técnico en la elaboración de planes de respuesta ante catástrofes naturales o provocadas, y fortaleciendo la resiliencia de sus sistemas agrícolas y alimentarios.

Del Marco estratégico 2010-2019 emana el Plan a Medio Plazo para 2014-2017, y el Programa de trabajo y presupuesto 2014-2015<sup>139</sup>.

### Gobernabilidad

Al tratarse de un organismo independiente, la FAO tiene constitución propia, firmada el 16 de Octubre de 1945. El Artículo 3 de la Constitución estipula que la "*Organización contará con una Conferencia, en la que cada Estado Miembro y Miembro Asociado estará representado por un delegado*"<sup>140</sup>(pág. 5).

- ✓ La Conferencia es el órgano rector supremo de la FAO. Se compone de 191 Miembros, 2 Miembros Asociados y una Organización Miembro. La Conferencia determina la política de la FAO, aprueba su presupuesto, y en general ejerce las facultades que le confiere su constitución. Se reúne de manera bianual.
- ✓ El Consejo está compuesto por 49 representantes de los Estados Miembros de los siete grupos regionales. Sirven por un periodo de tres años. Dentro de las atribuciones delegadas por la Conferencia, se reúne entre sus periodos de sesiones, generalmente tres veces a lo largo del bienio. Entre sus responsabilidades están la de la aprobación de nuevos miembros, y el presupuesto.
- ✓ Debajo del Consejo existen diferentes comités especializados, como el comité de programas y de finanzas, con labores principalmente de supervisión de diferentes funciones. Entre los principales, se encuentran el Comité de Programas, con la responsabilidad de revisar el programa de trabajo y presupuesto asociado de la FAO, el contenido de las actividades programáticas, y hacer recomendaciones sobre temas prioritarios. El Comité de Finanzas aprueba las transferencias presupuestarias y examina las cuentas auditadas<sup>141</sup>.

A nivel institucional, la FAO cuenta con siete departamentos: Agricultura y Protección del Consumidor, Pesca y Acuicultura, Desarrollo Económico y Social, Forestal, Servicios internos, Recursos Humanos, Finanzas y Cooperación Técnica. Estos departamentos tienen su sede central en Roma (Italia). A nivel descentralizado, la FAO

---

<sup>139</sup>FAO, C/2013/3

<sup>140</sup>FAO, Textos Fundamentales. Volumen I y II. Ed, 2013

<sup>141</sup> <http://www.fao.org/unfao/govbodies/gsbhome/es/>

cuenta con cinco oficinas regionales, 9 subregionales, 80 a nivel nacional, 3 con representación, y 38 cubiertos a través de acreditación conjunta.

### Recursos

El total del presupuesto de la FAO para 2014-2015 es de 2400 millones de USD. El 41% procede de cuotas provenientes de los Estados Miembros (1005,6 millones USD), mientras que el 59% se moviliza a través de contribuciones voluntarias de los Estados Miembros y otros organismos bilaterales y multilaterales (1394,4 millones USD)<sup>142</sup>.

A pesar de que el total del presupuesto no ha variado sustancialmente en los dos últimos bienios, se ha producido un incremento de la contribución voluntaria, y un descenso de las cuotas de los Estados Miembros (en el bienio 2008-2009 las cuotas regulares representaban el 48% del total del presupuesto<sup>143</sup>).

Las contribuciones de los Estados Miembros suponen el presupuesto ordinario, que se aprueba en la Conferencia bienal de la FAO. Las contribuciones voluntarias apoyan principalmente programas y proyectos de cooperación técnica, y otras actividades de asistencia técnica y de emergencia. Dichas actividades, a pesar de ser voluntarias deben tener objetivos claramente definidos, y alineados con el marco estratégico aprobado por la Conferencia de la FAO.

### Función de Evaluación

La evaluación es una función centralizada en la FAO. Existen oficiales de monitoreo en las oficinas regionales y en algunas nacionales, pero no tienen responsabilidades directamente vinculadas con la evaluación. La Oficina de Evaluación de la FAO existe desde 1968, por lo que se puede afirmar que está muy consolidada a nivel institucional.

Los procesos de reforma derivados del Plan de Acción Inmediato aprobado por la Conferencia en 2008, también tuvieron sus efectos en la función de evaluación. La Política de Evaluación, denominada Charter, se aprobó en Abril de 2010 por el Consejo, en alineamiento con las normas UNEG, y en claro entendimiento de la evaluación como parte del impulso de las reformas por la gestión por resultados en la FAO. Desde la aprobación de la Política, la Oficina de Evaluación es parte de las Oficinas del Director General. La Oficina de Evaluación reporta al Director General, aunque tiene canal directo de comunicación con el Consejo a través del Comité de Programas.

La Oficina de Evaluación es responsable por la realización de tres tipos de evaluaciones: evaluaciones para los órganos de gobierno (generalmente vinculadas a objetivos y áreas estratégicas del plan de trabajo), evaluaciones de país (sobre el total de las actividades implementadas por la FAO en un país concreto en un periodo concreto), y evaluaciones de proyecto (en muchos casos sobre proyectos estratégicos,

---

<sup>142</sup>FAO, CL 148/3

<sup>143</sup>FAO, C/2011/5A

de gran volumen o requeridas por los donantes). Cada uno de los tres tipos de evaluaciones siguen unos procedimientos de planificación, reporte y seguimiento diferentes.

La Oficina de Evaluación también es responsable de la preparación de informes a los órganos de gobierno, del seguimiento a las recomendaciones de los informes, y de la diseminación de resultados a través de actividades de comunicación<sup>144</sup>.

### Proceso de Revisión por pares

El proceso de revisión por pares se llevó a cabo entre Noviembre de 2011 y Noviembre 2012. La metodología utilizada incluyó una amplia revisión documental de la FAO, su estructura, y su función de evaluación. El panel se compuso de 4 miembros, incluyendo evaluadores/as profesionales vinculados a otras entidades de actividades relevantes para el mandato de FAO, como FMAM y PNUMA, y redes multilaterales (DAC Evaluation Network).

En Enero de 2012 tuvo lugar una fase de entrevistas en la sede de la FAO en Roma. Como parte del personal entrevistado se incluyó a personal de la Oficina de Evaluación, las Direcciones Técnicas, y miembros del Comité de Programas. También se analizó una muestra de cinco evaluaciones presentadas a los órganos de gobierno elaboradas entre 2009 y 201<sup>145</sup>. En Abril 2012 se efectuó una segunda fase de entrevistas en Roma, con el fin de clarificar aspectos hallados durante la revisión documental. El borrador del proceso de revisión fue presentado a la Oficina de Evaluación, y el Comité de Programas para comentarios preliminares, y el informe final se presentó en la reunión del Comité de Programas en Noviembre de 2012.

El informe de revisión por pares es públicamente accesible a través de la página web de la UNEG (UNEG/FAO 2012):

<http://www.uneval.org/document/detail/1191>

---

<sup>144</sup> Fuente: UNEG/FAO 2012 y <http://www.fao.org/evaluation/oed-home/en/>

<sup>145</sup> Las cinco evaluaciones fueron: 1) Strategic Evaluation of FAO Country Programming, including the NMTPF mechanism, 2) Evaluation of FAO's role and work related to water, 3) Evaluation of FAO's role and work in nutrition, 4) Evaluation of FAO's regional and sub-regional offices for the Near East, y 5) Second Real Time evaluation of FAO's work on highly pathogenic avian influenza.

## 1.7 FMAM

### Mandato

El FMAM se estableció en 1991 como un programa piloto del Banco Mundial con el fin de asistir a los países en desarrollo y economías en transición a cumplir con las obligaciones derivadas de acuerdos multilaterales en materia medioambiental, y favorecer la armonización entre políticas nacionales, regionales y globales<sup>146</sup>.

El PNUD, PNUMA y el Banco Mundial fueron denominadas en un inicio como las tres ejecutoras de los proyectos financiados con fondos FMAM. Tres años más tarde, en 1994, el FMAM se convirtió en organismo independiente tras la Cumbre para la Tierra en 1992. Desde entonces, el Banco Mundial actúa como entidad depositaria del FMAM y le proporciona servicios administrativos.

A mayo de 2015, el FMAM tiene un total de 183 miembros, y es el mecanismo financiero designado para la implementación de los siguientes convenios y convenciones internacionales<sup>147</sup>:

- ✓ Convenio sobre la Diversidad Biológica (CBD)
- ✓ Convención Marco de NNUU sobre el Cambio Climático (CMNUCC)
- ✓ Convenio de Estocolmo sobre los Contaminantes Orgánicos Persistentes (POPs)
- ✓ Convención de NNUU para Combatir la Desertificación (CNULD)
- ✓ Convenio de Minimata sobre el Mercurio

El FMAM otorga contribuciones para financiar proyectos en las áreas prioritarias cubiertas por las convenciones y los convenios internacionales mencionados como biodiversidad, cambio climático, aguas internacionales, degradación de la tierra, la capa de ozono y contaminantes orgánicos persistentes. Para ello, FMAM ofrece cofinanciamiento en condiciones concesionarias para cubrir lo que se denominan costos incrementales, que son aquellos costos adicionales para poder favorecer que las iniciativas a nivel nacional proporcionen beneficios ambientales a nivel mundial.

Como fondo, su naturaleza es puramente financiera, no implementa proyectos de por sí. Los fondos se otorgan a través de ciclos de cuatro años a diferentes agencias<sup>148</sup>. Estas agencias son responsables de desarrollar los documentos de proyecto y gestionar la implementación de los proyectos una vez aprobados. Actualmente hay 14 agencias implementadores de fondos FMAM:

- ✓ El Programa de las Naciones Unidas para el Desarrollo (PNUD)

---

<sup>146</sup><https://www.thegef.org/gef/whatisgef>

<sup>147</sup>Además de los convenios y convenciones citadas, FMAM respalda la aplicación del Protocolo de Montreal relativo a las Sustancias que Agotan la Capa de Ozono en países con economías en transición.

<sup>148</sup>En un principio se estableció una diferencia entre agencias implementadoras y ejecutoras. Con el tiempo esas diferencias se eliminaron y ahora son todas igualmente referidas como agencias FMAM.

- ✓ El Programa de las Naciones Unidas para el Medio Ambiente (PNUMA)
- ✓ El Banco Internacional para la Reconstrucción y Desarrollo (BM-BIRF)
- ✓ La Organización Mundial de las Naciones Unidas para la Alimentación y la Agricultura (FAO)
- ✓ La Organización de las Naciones Unidas para el Desarrollo Industrial (ONUDI)
- ✓ El Banco Africano de Desarrollo
- ✓ El Banco Asiático de Desarrollo
- ✓ El Banco Europeo para la Reconstrucción y el Desarrollo
- ✓ El Banco Interamericano de Desarrollo
- ✓ El Fondo Internacional de Desarrollo Agrícola (FIDA)
- ✓ Conservación Internacional
- ✓ El Banco de Desarrollo de Sudáfrica
- ✓ La Unión Internacional para la conservación de la naturaleza
- ✓ El Fondo Mundial para la Vida Salvaje

### Marco Estratégico

Los recursos del FMAM se renuevan cada cuatro años, a través de lo que se denominan procesos de reposición de fondos. Para ello, las negociaciones que abarca el proceso de reposición incluyen la discusión del marco programático (Documento de Programa), y la revisión del Estudio de Rendimiento General (Overall Performance Studies), que es realizada por la Oficina de Evaluación sobre los proyectos financiados por el FMAM durante el periodo de reposición anterior.

Actualmente el FMAM se encuentra financiando la implementación de la sexta reposición de fondos. El Documento de Programa FMAM-6 contiene las líneas programáticas durante el ciclo Junio 2014 a Junio 2018. En dicho documento se señalan 5 áreas prioritarias de acción del FMAM<sup>149</sup>:

- ✓ Biodiversidad. Esta área prioritaria se enfoca en mantener la biodiversidad y los servicios que el ecosistema proporciona. Con el fin de lograr este objetivo, la estrategia está orientada a: Mejorar la sostenibilidad de los sistemas de áreas protegidas; Reducir las amenazas a la biodiversidad; El uso sostenible de la biodiversidad; Manejo y gestión integrada de la biodiversidad en los sectores productivos. La estrategia de biodiversidad de FMAM-6 se compone de un total de 10 programas, donde cada uno de ellos pretende dar respuesta a las amenazas y oportunidades temáticamente cubiertas por esta área. Además, el programa contempla uno de los factores encubiertos en la pérdida de biodiversidad: La falta de rendición de cuentas y el valor económico global de los ecosistemas y la biodiversidad.
- ✓ Mitigación de cambio climático. Se pretende prestar apoyo a los países en desarrollo y economías en transición a realizar cambios orientados a una reducción en sus emisiones de efecto invernadero. Esta área temática tiene tres

---

<sup>149</sup>GEF 6-Programming Directions, May 2014

objetivos fundamentales: Promover la innovación y transferencia de tecnología; Demostración de opciones innovadoras de mitigación con posibilidad de impacto y replicabilidad; Crear condiciones para poner en marcha políticas de mitigación como parte de las estrategias de desarrollo sostenibles. Esta estrategia se compone de cinco programas que apoyan los tres objetivos mencionados. Estos programas representan una serie de medidas para gestionar los riesgos todavía existentes en la transformación hacia un camino de bajas emisiones.

- ✓ Productos químicos y residuos. En el largo plazo, esta área pretende reducir la exposición humana y medioambiental a productos químicos y residuos tóxicos incluyendo POPs, mercurio y otras sustancias a través de la reducción significativa de la producción, uso, consumo y emisiones de dichos productos químicos y residuos. Para ello, la estrategia incluida en FMAM-6 trata de combinar la puesta en marcha de tecnologías y sistemas seguros ambientalmente con mecanismos financieros y organizacionales, políticas y prácticas que faciliten ayuda a los países hacia ese cambio necesario transformativo. Esta área se compone de seis programas. Sin embargo, la enorme complementariedad de estos programas hacen que sea posible la financiación de un enfoque integrado que incluya varios de estos programas.
- ✓ Aguas internacionales. El objetivo a largo plazo de esta área prioritaria es el de la promoción de una gestión colectiva e integrada de los sistemas de agua inter fronterizos. Para ello, a través de proyectos financiados por FMAM, se facilitará la cooperación y diseminación de conocimiento entre estados a través de capacitación, investigación orientada, y diseminación de buenas prácticas; la armonización y coordinación de la gestión de los sistemas de aguas transfronterizos superficiales con aguas en profundidad; y la promoción y catalización de inversiones para promover sistemas de pesca sostenibles, proteger o restaurar hábitats costeros y reducir la polución en las costas.
- ✓ Degradación de la tierra. El principal objetivo de esta área temática es hacer frente y revertir los procesos actuales de degradación de la tierra, especialmente la desertificación y deforestación, a través de la promoción de buenas prácticas que desde lo local puedan también proporcionar impacto a nivel global. A través de FMAM-6, se apoyan oportunidades de gestión pecuaria y reforestación de zonas rurales, y se promueve una gestión integrada de producción que fortalezca los niveles de resiliencia en los ecosistemas agropecuarios.

### Gobernabilidad

En cuanto a su estructura de gobierno, el FMAM se compone de los siguientes órganos:

- ✓ El órgano supremo del FMAM es la Asamblea donde están representados cada uno de los países participantes. La Asamblea se reúne cada 3 años. Es

responsable de revisar y evaluar las políticas generales de FMAM, sus operaciones, y aprobar la membresía de nuevos participantes.

- ✓ La Asamblea delega en el Consejo del FMAM las decisiones relativas al trabajo del FMAM entre sus sesiones. El Consejo consiste en 32 miembros representantes de países en desarrollo y países desarrollados elegidos para periodos de 3 años con posibilidad de reelección. El Consejo se reúne dos veces al año en la Secretaría del FMAM en Washington. Las reuniones son presididas por el Oficial Ejecutivo del FMAM y un co-presidente elegido entre los miembros del Consejo.
- ✓ La Secretaría del FMAM tiene un/a máximo/a responsable que es el/la Oficial Ejecutivo. Es nominado/a por el Consejo a través de las recomendaciones de las agencias implementadoras del FMAM. La Secretaría reporta al Consejo y la Asamblea. Entre sus funciones se incluye la elaboración del plan anual de trabajo, y la puesta en marcha de medidas en la organización para asegurar la operatividad de las políticas y estrategias. A pesar de que el FMAM es una entidad independiente, la Secretaría es administrativamente parte del Banco Mundial. La Oficina de Evaluación es completamente independiente de la Secretaría del FMAM. El/la Directora/a de la Oficina de Evaluación reporta directamente al Consejo, al igual que el Oficial Ejecutivo del FMAM.

## Recursos

Originalmente en 1991 el FMAM fue dotado de 1000 millones de USD. Se calcula que desde su creación en 1991, el FMAM ha otorgado cerca de 13.500 Millones USD, y ha co-financiado cerca de 4000 proyectos por un valor total de 65.000 Millones USD<sup>150</sup>.

En total el FMAM se ha repuesto de fondos en seis ocasiones, con los siguientes importes:

FMAM 1 (1994-1998): 2.000 Millones USD  
FMAM 2 (1998-2002): 2075 Millones USD  
FMAM 3 (2002-2006): 3000 Millones USD  
FMAM 4 (2006-2010): 3130 Millones USD  
FMAM 5 (2010-2014): 4340 Millones USD  
FMAM 6 (2014-2018): 4433 Millones USD

Los donantes principales son Estados Unidos, Japón, Alemania, Reino Unido y Francia. La lista de donantes incluye otros países como China, India, Méjico, Nigeria, y Sudáfrica.

Las actividades y recursos necesarios para mantener la Oficina de Evaluación forman parte del Programa y presupuesto aprobado por el Consejo. En este sentido, un total

---

<sup>150</sup><https://www.thegef.org/gef/whatisgef>

de 19 Millones de USD fueron aprobados para el funcionamiento de la Oficina de Evaluación para el periodo 2014-2018<sup>151</sup>.

### Función de Evaluación

La Oficina de Evaluación se estableció en 2003 fruto de las negociaciones para la tercera reposición de fondos (*Third replenishment*). Como se ha mencionado anteriormente, la Dirección de la Oficina de Evaluación reporta directamente al Consejo y es nombrada por el Consejo por un periodo renovable de 5 años.

En el año 2006 el Consejo aprobó la Política del FMAM sobre Monitoreo y Evaluación, que fue actualizada posteriormente en 2010. Desde su aprobación, la Oficina de Evaluación ha formulado diferentes guías y procedimientos para facilitar la implementación de la Política de Evaluación<sup>152</sup>.

Las funciones de la Oficina de Evaluación son diversas y entre ellas destacan las siguientes: Realizar evaluaciones de programas y actividades en los países receptores de fondos FMAM; establecer normas, procedimientos y metodologías para armonizar la aplicación de actividades de monitoreo y evaluación en agencias FMAM; poner en marcha controles de calidad para asegurar que la evaluación efectuada desde las agencias FMAM responde a estándares de calidad; y disseminar la información proveniente de evaluaciones.

Entre los tipos de documentos que la Oficina de Evaluación produce, además de las guías y procedimientos metodológicos, se incluyen los estudios de rendimiento general, evaluaciones de programas, evaluaciones de actividades de país e informes anuales de rendimiento.

### Proceso de revisión por pares

Cada cuatro años el FMAM produce lo que se denomina el Estudio de Rendimiento General (Overall Performance Study, OPS), como instrumento base en la deliberación de los proceso de reposición de fondos.

En Junio de 2007, la Oficina de Evaluación del FMAM propuso ser la responsable de la realización del estudio OPS, como parte de su plan de trabajo. El Consejo aprobó la propuesta, con la condición de que las partes del estudio que se refieren a monitoreo y evaluación no formaran parte del análisis, para evitar un posible conflicto de interés con el trabajo de la Oficina de Evaluación. En su lugar, se recomendó que la función de evaluación en el FMAM fuera analizada por un proceso independiente de revisión por pares.

---

<sup>151</sup>GEF 6-Programming Directions, May 2014

<sup>152</sup>Entre ellas se incluyen: Ethical guidelines of the Evaluation Office, Measuring results through SMART way, Guidelines to conduct terminal evaluations, Terms of Reference for Country Portfolio Evaluations

En consecuencia, el informe del proceso de revisión por pares del FMAM fue realizado y presentado como parte integrante de las reuniones de deliberación de la quinta reposición de fondos del FMAM (que tuvieron lugar de Octubre 2009 a Mayo de 2010), y como complemento al informe de Estudio de Rendimiento General realizado por la Oficina de Evaluación del FMAM sobre los resultados de la cuarta reposición de fondos.

El proceso de realización del informe de revisión por pares tuvo una duración de alrededor de 1 año, de Mayo de 2008 a Mayo de 2009.

El panel estuvo compuesto por seis integrantes, evaluadores/as profesionales provenientes de otros organismos de NU, como el Programa Mundial de Alimentos (PMA), cooperación bilateral de desarrollo (Bélgica y Finlandia), y organizaciones multilaterales (Eurasian Development Bank, Organización Internacional para la Cooperación en Evaluación).

Entre la metodología aplicada para el proceso de revisión por pares se incluyó:

- ✓ Extensa revisión documental sobre el FMAM, su estructura, su función de evaluación, presupuesto y plan de trabajo, y las deliberaciones del Consejo en torno al establecimiento y seguimiento de la evaluación.
- ✓ Entrevistas con miembros del Consejo del FMAM, personal de la Secretaría del FMAM, personal de la Oficina de Evaluación, de agencias del FMAM y puntos focales del FMAM en los países. Estas entrevistas se realizaron durante las visitas a Washington (a la sede del FMAM, el Banco Mundial y el Banco Interamericano de Desarrollo), a Nueva York (PNUD), Filipinas (Banco Asiático de Desarrollo), Kenia (PNUMA) y Roma (FAO y FIDA). Las entrevistas a personal de la ONUDI se realizaron durante la reunión del Consejo del FMAM in Octubre de 2008. En total se realizaron cerca de 100 entrevistas.

Para el análisis de la calidad y contenido de los informes se tomó una muestra de siete informes de evaluación<sup>153</sup>. Los miembros de panel de revisión por pares tuvieron oportunidad de realizar diferentes reuniones para la preparación de la revisión, armonizar la metodología, intercambiar información sobre los hallazgos, y discutir el informe de revisión.

El informe de revisión por pares es accesible a través de la página web de UNEG (UNEG/GEF 2009):

<http://www.uneval.org/document/detail/313>

---

<sup>153</sup> Los informes fueron: the Philippines Country Portfolio Evaluation, Kenya country reports, the Annual Report on Impact 2007, the RAF Mid-Term Review, the Joint Evaluation of the GEF Activity Cycle and modalities, the Role of Local Benefits in Global Environmental Programs, and the 2007 Annual Performance Report.

Posteriormente, en el periodo entre Enero y Junio de 2014, se ha realizado una segunda revisión por pares de la función de evaluación del FMAM, con el objetivo de dar seguimiento a las recomendaciones emanadas del anterior informe, y enfocarse en detalle en algunas de las debilidades encontradas en la primera revisión por pares. La metodología fue muy similar al proceso de revisión anterior, e incluyó una extensa revisión documental, una visita a la Secretaría del FMAM (del 21 al 24 de Marzo de 2014), y entrevistas con el personal de la Oficina de Evaluación, miembros del Consejo, y puntos focales de agencias implementadoras.

El informe de la segunda revisión es accesible en la página web de la UNEG (UNEG/GEF 2014):

<http://www.uneval.org/document/detail/1624>

## **1.8 OIOS**

### Mandato

OIOS fue establecida por la Asamblea de NU a través de la resolución 48/218B de 1994 con el propósito de asistir a la Secretaría General de NU en sus funciones de supervisión y control de los recursos. Para ello, OIOS se estableció como una oficina independiente con diferentes áreas de control: auditoría interna, monitoreo, inspección, evaluación y servicios de investigación.

Funcionalmente, OIOS está ubicada por debajo de la Secretaría General de NU, y presta servicios de supervisión sobre todas las actividades de Naciones Unidas bajo la responsabilidad del Secretario/a General. Esto abarca las secretarías de NU en Nueva York, Nairobi, Viena y Ginebra, las comisiones regionales, operaciones de paz y humanitarias, y asistencia a los Fondos y los Programas administrados bajo la autoridad de la Secretaría General (PNUMA, UNO-Habitat, UNHCR, y OHCHR).

La labor de OIOS consiste principalmente en asegurar que las actividades realizadas están de acuerdo a los programas aprobados, las regulaciones existentes, y recomendar medidas para una implementación eficiente y efectiva. También, en prevenir y detectar el fraude, mal uso, abuso o mala gestión de los recursos. Y contribuir a detectar todas aquellas buenas prácticas y lecciones aprendidas orientadas a mejorar el rendimiento y la efectividad de los programas.

### Marco Estratégico

Puesto que se trata de un organismo de supervisión de actividades ejecutadas y financiadas por otras entidades, su estrategia pasa por promover un sistema de rendición de cuentas que sea transparente y efectivo, y en fortalecer las capacidades de la Secretaría para identificar y mitigar riesgos que puedan dificultar el logro de objetivos y resultados. En particular, OIOS:

- ✓ Proporciona información y análisis independientes que faciliten la toma de decisiones
- ✓ Realiza análisis independientes de la efectividad en el uso de los recursos
- ✓ Propone medidas para asistir a las diferentes entidades, fondos y programas en cómo responder a riesgos y beneficiarse de oportunidades
- ✓ Promueve una cultura de cambio, orientado a resultados, y gestión efectiva.

### Gobernabilidad

OIOS está dividida en 3 secciones: Auditoría Interna, Investigaciones y Evaluación e Inspecciones. El/la responsable de la OIOS tiene rango de Sub-Secretario/a General, y reporta directamente al Secretario/a General. Las actividades llevadas a cabo por OIOS son llevadas a la consideración de la Asamblea General (Quinto Comité) a través del Secretario/a General.

En NU existen otras entidades también encargadas de la supervisión y el control. En este sentido, OIOS coordina sus actividades con la Unidad de Inspección Conjunta, y la Junta de Auditores a través de reuniones regulares.

## Recursos

Los recursos aprobados por la Asamblea General para OIOS en el bienio 2014-2015 ascienden a 121,000 Millones USD<sup>154</sup>. Por secciones, el departamento de evaluaciones e inspecciones tiene asignado un 20% del presupuesto, frente al 38% que va destinado a auditoría interna, y 27% a investigaciones. El restante 15% se destina a gastos de dirección y apoyo a programas. Existe así un alto nivel de detalle y transparencia en relación a los recursos asignados para cada actividad de OIOS, puesto que en su mayor parte están incluidos en el presupuesto ordinario de la Secretaría.

## Función de Evaluación

En cuanto a la función de evaluación específicamente, ésta se estableció formalmente el 1 de enero de 2008 en seguimiento a los acuerdos adoptados entre los estados miembros y la Secretaría de las NU en la Cumbre Mundial de 2005<sup>155</sup> sobre la reforma de la gestión de las NU y, en particular, el fortalecimiento de su capacidad de supervisión. Su mandato queda establecido en el Boletín del Secretario General ST/SGB/2000/8<sup>156</sup> en el que se establece que *“la función de evaluación debe contribuir a determinar de manera sistemática y objetiva la relevancia, eficiencia, efectividad e impacto de las actividades de un organismo determinado en relación con sus objetivos, con el propósito de incrementar la efectividad de los programas modificando su contenido, y si es necesario, revisando sus objetivos”* (ST/SGB/2000/8, pág. 12).

La sección de evaluación e inspecciones adquiere así un papel ciertamente sustantivo dentro de OIOS, e inicia, a partir de su integración formal en 2008, el desarrollo de evaluaciones programáticas así como evaluaciones temáticas e inspecciones generalmente sobre procesos específicos y/o unidades dentro de un organismo determinado. A su vez, esta sección dentro de OIOS realiza de manera bianual el informe denominado *“Fortalecer el rol de evaluación y la aplicación de los hallazgos evaluativos en el diseño de programas, asistencia y las directivas políticas”*.

El plan de evaluación de OIOS es preparado en base a un criterio de riesgo. Este criterio asegura transparencia en la selección de áreas u organismos, a la vez que relevancia y puntualidad de las evaluaciones. Es aprobado por el Quinto Comité de la Asamblea General.

---

<sup>154</sup>A/68/6

<sup>155</sup> Ver [http://www.un.org/spanish/summit2005/fact\\_sheet.html](http://www.un.org/spanish/summit2005/fact_sheet.html)

<sup>156</sup> Accesible en

<http://daccess-dds-ny.un.org/doc/UNDOC/GEN/N00/408/48/PDF/N0040848.pdf?OpenElement>

### Proceso de revisión por pares

El proceso de revisión por pares se extendió por alrededor de 1 año, de Enero de 2008 a Enero de 2009. Durante ese periodo, el panel, compuesto por 6 miembros procedentes de otras agencias de NU (FAO, UNO Mujeres), cooperación bilateral para el desarrollo (Noruega y Suiza), se centró en el análisis de la función de evaluación desde los tres criterios fundamentales aprobados en la metodología referente al proceso de revisión por pares en NU: utilidad, independencia y credibilidad.

La metodología incluyó una extensa revisión documental (incluyendo informes realizados por OIOS, informes de la Asamblea General, y resoluciones de NNUU entre otros vinculados al trabajo y operatividad de OIOS), y dos fases de entrevistas en la Sede en Nueva York que incluyeron a personal de la Oficina de OIOS, el Presidente y Vicepresidente del Comité Asesor de Auditoría Independiente y miembros del Quinto Comité y Comité de Programas y Coordinación de la Asamblea General de NNUU.

El informe borrador fue presentado y discutido con el personal de la Sección de Evaluación e Inspecciones de OIOS. Posteriormente el informe final fue formalmente presentado a OIOS en Marzo de 2009.

El informe es públicamente accesible a través de la página web de UNEG (UNEG/OIOS 2009):

<http://www.uneval.org/document/detail/226>

## 1.9 PMA

### Mandato

El Programa Mundial de Alimentos (PMA) fue establecido en 1961 a través de resoluciones paralelas en ECOSOC y FAO. Su mandato es erradicar el hambre en el mundo. En un principio se estableció como un proyecto piloto que debería entrar en funcionamiento en 1963. Sin embargo, las catástrofes y conflictos del terremoto en Irán en Septiembre de 1962, el huracán en Tailandia en Octubre de ese mismo año, y los cerca de 5 millones de refugiados tras la independencia de Argelia hizo absolutamente necesaria la ayuda del PMA, y ha continuado sus operaciones desde entonces.

A pesar de que su principal función es la de proveer asistencia para la seguridad alimentaria en situaciones de emergencia, con el tiempo, su ámbito de actuación, ha evolucionado hacia una perspectiva integrada de la ayuda alimentaria. Así, el PMA también se encarga de aspectos relacionados con la prevención, mitigación de riesgos y respuestas sostenibles a las crisis alimentarias. De la misma manera, con el tiempo se ha convertido en la mayor agencia de ayuda humanitaria dentro de NU. Desde el 2003 incluye entre sus responsabilidades la provisión de servicios de transporte aéreo para toda la ayuda humanitaria coordinada o gestionada en el marco de NU.

Desde el 2007, el PMA es responsable de la coordinación de la Red de Respuesta Humanitaria (Humanitarian Response Depots)<sup>157</sup>, con el fin de favorecer las tareas de coordinación y mejorar la eficiencia de la asistencia, a través del uso de puntos estratégicos de reservas y abastecimiento. El PMA, participa en el Comité de Seguridad Alimentaria Mundial<sup>158</sup>, contribuye al Marco Estratégico Mundial para la Seguridad Alimentaria y la Nutrición, y al programa de desarrollo sostenible para después de 2015.

### Marco Estratégico

La dirección estratégica del organismo está representada en su Plan Estratégico, el cual se renueva cada cuatro años. Actualmente está en vigor el Plan Estratégico 2014-2017, en donde se establece el marco de prioridades en la prestación de asistencia alimentaria para las personas más vulnerables y con falta de recursos. El Plan Estratégico detalla la actuación del PMA para contribuir a la consecución de las metas globales más amplias: reducir el riesgo y la vulnerabilidad a las crisis, romper el ciclo del hambre, y lograr una situación sostenible en cuanto a seguridad alimentaria y nutrición, con arreglo a lo dispuesto en la iniciativa 'Reto del hambre cero' del Secretario General de las Naciones Unidas (Plan Estratégico PMA 2014-2017, pág. 3).

Las cuatro prioridades del Plan Estratégico 2014-2017<sup>159</sup> son:

---

<sup>157</sup><http://www.wfp.org/logistics/humanitarian-response-depot>

<sup>158</sup><http://www.fao.org/cfs/cfs-home/es/>

<sup>159</sup> WFP/EB.A/2013/5-A/1

- ✓ Fortalecimiento y protección de los medios de subsistencia en las emergencias. La intervención en situaciones de emergencia humanitaria es una prioridad fundamental para el PMA. Para ello, en colaboración con ACNUR, UNICEF, FAO, OMS y otras organizaciones con mandatos relevantes en ayuda humanitaria, el PMA ayuda a los gobiernos a proteger los medios de subsistencia en las emergencias, y a facilitar medidas para atender las necesidades alimentarias y nutricionales urgentes. Esto incluye mantener estrecha colaboración con los gobiernos, comunidades y organizaciones en la evaluación de las necesidades, en la puesta en marcha de medidas y respuestas apropiadas en el contexto de cada país.
- ✓ Respaldo o restablecer la seguridad alimentaria y la nutrición. Esta área prioritaria está destinada a los países que salen de situaciones de inestabilidad o de conflictos, y donde es posible que, ante infraestructuras sociales y económicas precarias o poco equitativas, puedan darse situaciones de inseguridad alimentaria. El PMA colabora con los gobiernos y otras organizaciones internacionales en el apoyo a programas que respondan a las necesidades alimentarias y nutricionales de los grupos más vulnerables, incluyendo mujeres, niños/as, personas refugiadas, y apoya a las comunidades en zonas de reasentamiento a poner en marcha medidas para hacer frente a la inseguridad alimentaria de forma sostenible.
- ✓ Reducir los riesgos, y facilitar ayuda para que los países puedan satisfacer sus propias necesidades alimentarias y nutricionales. A través de esta área prioritaria, el PMA colabora con los gobiernos y otros organismos internacionales, como FAO, y el FIDA, la sociedad civil y el sector privado para ayudar a prevenir, gestionar y reducir los riesgos, fomentar la resiliencia y disminuir la dependencia externa. Esto incluye apoyar y facilitar el acceso de los pequeños agricultores a los mercados, apoyar el empoderamiento económico de grupos vulnerables, y prestar asistencia en la búsqueda de opciones auto sostenibles para satisfacer necesidades alimentarias y nutricionales.
- ✓ Reducir la desnutrición y romper el ciclo intergeneracional del hambre. Desde esta área prioritaria se apoya, en colaboración con ONUSIDA, FAO, UNICEF y la OMS, a los gobiernos en sus programas de reducir la desnutrición crónica, y poner en práctica programas de nutrición, en alineamiento con las políticas nacionales y las metas mundiales. Para ello, actúa en el marco de plataformas multisectoriales, como el movimiento SUN (para el fomento de la nutrición), con el fin de desarrollar y aplicar medidas eficaces y sostenibles. Particular atención se presta a las necesidades nutricionales de los grupos vulnerables, incluyendo mujeres, niños/as, y las personas afectadas por el VIH, la tuberculosis y la malaria.

## Gobernabilidad

La Junta Ejecutiva es el órgano rector de las actividades de PMA, tal y como está estipulado en la resolución 48/162 de la Asamblea General de NU. Está formado por 36 miembros, de los cuales 18 son elegidos por el Consejo Económico y Social de las NNUU (ECOSOC), y 18 por el Consejo de la FAO. Se reúne 3 veces al año en la sede de PMA en Roma (Italia)<sup>160</sup>.

Las funciones de la Junta Ejecutiva incluyen:

- ✓ Coordinar y dirigir las políticas de ayuda alimentaria de corto y largo plazo
- ✓ Supervisar la gestión de PMA
- ✓ Revisar, y aprobar los programas, proyectos y actividades, presupuestos, y en general la ejecución de los programas.
- ✓ Informar anualmente al ECOSOC y al Consejo de FAO sobre las actividades implementadas en el marco de trabajo de PMA.

El/la Director/a Ejecutivo/a es el máximo responsable de las operaciones en PMA, es elegido/a por la Junta Ejecutiva por periodos de 5 años.

La gran mayoría del personal profesional trabaja desde las oficinas de campo (90 países). La sede en Roma está estructurada en cinco departamentos operativos, que incluyen la programación de asistencia alimentaria, logística, movilización de recursos, comunicación y supervisión (auditoría y evaluación).

## Recursos

El 100% de las contribuciones al PMA son voluntarias. Las contribuciones pueden ser en dinero, alimentos, o utensilios básicos para plantar, almacenar o cocinar alimentos.

La mayoría de las contribuciones son gubernamentales (alrededor de 60), seguido de grandes corporaciones, y donaciones individuales.

El total de contribuciones al PMA a Mayo de 2015 asciende a US\$ 1,956,023,878. La lista de los diez países con mayores contribuciones incluye Estados Unidos, Reino Unido, Canadá, Japón, Holanda, la Unión Europea, Suiza, UN CERF, Alemania y Noruega<sup>161</sup>.

## Función de Evaluación

La función de evaluación existe en PMA desde 1965, y es una función mixta compartida entre las evaluaciones corporativas que se hacen desde la sede por la

---

<sup>160</sup><http://www.wfp.org/about/corporate-information>

<sup>161</sup><http://www.wfp.org/about/funding/year/2015>

Oficina de Evaluación, y las evaluaciones que se hacen desde las oficinas de campo y las sedes regionales.

La Política de Evaluación fue aprobada por la Junta Ejecutiva en 2008, y marca las líneas directrices de la función de evaluación. Su establecimiento ha servido para impulsar la aplicación de otras prácticas evaluativas relacionadas, como el sistema de seguimiento a las recomendaciones y un sistema de control de calidad de las evaluaciones. Las evaluaciones corporativas son accesibles públicamente, y las realizadas a nivel descentralizado son resumidas en el Informe Anual que presenta el/a Director/a de Evaluación a la Junta Ejecutiva<sup>162</sup>.

El/a Director/a de Evaluación reporta directamente al Director/a Ejecutivo/a del PMA. El Plan de trabajo de la Oficina de Evaluación forma parte del plan trianual gerencial aprobado por la Junta Ejecutiva.

Los diferentes tipos de evaluaciones que se llevan a cabo en PMA son:

- ✓ Evaluaciones estratégicas
- ✓ Evaluaciones de políticas
- ✓ Evaluaciones de portafolio de país
- ✓ Evaluaciones de impacto
- ✓ Evaluaciones de procedimientos operativos

### Proceso de revisión por pares

El PMA ha realizado dos procesos de revisión por pares de su función de evaluación, el primero en 2007 y posteriormente en 2013.

El primer proceso de revisión por pares se llevó a cabo por cinco miembros de panel, evaluadores/as profesionales provenientes de otras agencias de NU (OCHA), cooperación bilateral (Suiza y Holanda), organismos/redes internacionales (EVALNET), y ONGs (CARE).

La metodología aplicada incluyó una profunda revisión documental, cuestionarios (enviados a más de 480 personas), entrevistas a directores/as regionales, y de las oficinas centrales, personal de la unidad de evaluación (un total de 124 entrevistas), y un ejercicio de meta-evaluación para analizar la calidad y nivel de estandarización de los informes de evaluación, con una composición mixta de evaluaciones realizadas desde la Oficina de Evaluación y las oficinas regionales/nacionales.

La información se complementó con visitas de campo a las Oficinas de Johannesburgo, Yakarta, Bangkok, Vientián y Lilongüe, en función de la muestra de evaluaciones seleccionadas y respetando la cobertura geográfica de las actividades de PMA. Por último, se procedió a una reunión grupal con 14 miembros de la Junta Ejecutiva. El

---

<sup>162</sup>Fuentes: UNEG/PMA 2007, 2013 y <http://www.wfp.org/evaluation>

borrador fue discutido con PMA en Octubre de 2007, y la versión final se presentó oficialmente a la Oficina de Evaluación en Noviembre de 2007.

Los datos utilizados para el análisis de QCA responden a los resultados de esta primera revisión por pares. El informe es públicamente accesible en la web de UNEG (UNEG/PMA 2007):

<http://www.uneval.org/document/detail/97>

El segundo proceso de revisión por pares se llevó a cabo en 2013. Una de los principales propósitos de este segundo proceso fue analizar la evolución de la función de evaluación a raíz de las recomendaciones del primer informe, y la consecuente aprobación de la Política de Evaluación por la Junta Ejecutiva en 2008. En esta ocasión el panel estuvo compuesto por tres miembros, de los cuales uno era miembro del panel de la anterior revisión (OCHA), y los otros dos provenían de UNICEF y la cooperación bilateral holandesa.

El proceso total se extendió por un año, de Enero de 2013 a Febrero de 2014. La metodología fue similar al proceso anterior, comenzando por una revisión documental, una muestra de 20 informes de evaluación para proceder al examen de su calidad, 2 visitas a la sede en Roma donde tuvieron oportunidad de entrevistar a directores, miembros de la Junta Ejecutiva y personal de la Oficina de Evaluación, entrevistas telefónicas a directores regionales y de las oficinas nacionales y cuestionarios en línea (para el personal de PMA, y para los miembros de la Junta Ejecutiva).

El informe está disponible en la página de UNEG (UNEG/PMA 2014):

<http://www.uneval.org/document/detail/1623>

## 1.10 ONU Mujeres

### Mandato

ONU Mujeres fue establecida por la Asamblea General de las NU en Julio de 2010 a través de la resolución 64/289 con el fin de proporcionar, *"mediante sus funciones de apoyo normativo y actividades operacionales, asesoramiento y apoyo técnico a todos los Estados Miembros, de todos los niveles de desarrollo y de todas las regiones, a solicitud de éstos, sobre la igualdad entre los géneros, el empoderamiento y los derechos de la mujer y la incorporación de la perspectiva de género"* (A/RES/64/289 pág. 10).

Su establecimiento supone la consolidación y transferencia de algunos organismos y entidades existentes, como la Oficina del Asesor Especial en Cuestiones de Género, la División para el Adelanto de la Mujer de la Secretaría, el Fondo de Desarrollo de las Naciones Unidas para la Mujer y el Instituto Internacional de Investigaciones y Capacitación para la Promoción de la Mujer, para que funcione como secretaria y también desempeñe actividades operacionales en los países con la función principal de promover la igualdad y el empoderamiento de las mujeres.

Sus principales funciones incluyen: proporcionar apoyo técnico a entidades intergubernamentales en la formulación de políticas y estándares a nivel mundial; asistir a los Estados Miembros en la correcta aplicación de las políticas en plena coordinación con la sociedad civil; liderar y coordinar las actividades que se realicen desde el marco de NU sobre igualdad de género, y dar seguimiento a los progresos y dificultades en el desarrollo de las intervenciones ejecutadas en el sistema de NU<sup>163</sup>.

### Marco Estratégico

Los Planes Estratégicos en ONU Mujeres son aprobados por periodos de 3 años por la Junta Ejecutiva. Actualmente está en vigor el Plan Estratégico 2014-2017. En él se desglosan 6 principales áreas de actuación donde ONU Mujeres se compromete a contribuir activamente teniendo en cuenta su mandato, experiencia y ventaja comparativa. Las áreas son las siguientes<sup>164</sup>:

- ✓ Empoderamiento y participación en la toma de decisiones. Para ello, ONU Mujeres apoya la adopción y aplicación de reformas constitucionales, leyes y políticas que promueven la participación de la mujer; la gestión electoral con una perspectiva de género; y la creación de capacidad y cambios institucionales para atraer a mujeres hacia puestos de liderazgo en las instituciones de prestación de servicios.

---

<sup>163</sup>El Marco rector de las funciones de ONU Mujeres se encuentra en: La Convención sobre la Eliminación de Todas las Formas de Discriminación Contra La Mujer y el Protocolo Facultativo; la Declaración y Plataforma de Acción de Beijing; La Resolución 1325 del Consejo de Seguridad de la ONU sobre Mujeres, Paz y Seguridad; y la Declaración del Milenio y los Objetivos de Desarrollo del Milenio.

<sup>164</sup>UNW/2013/6

- ✓ Empoderamiento económico para las mujeres más pobres y en situación de exclusión. En esta área estratégica, ONU Mujeres presta especial atención a la adaptación y aplicación de leyes y políticas para fortalecer el empoderamiento de la mujer; el establecimiento y puesta en marcha de servicios que tienen en cuenta la perspectiva de género, incluidos el transporte, los mercados, servicios energéticos para fomentar los medios de vida sostenibles de la mujer; el apoyo a la promoción de la igualdad de género, para conseguir que tenga una mayor influencia en las políticas económicas y las estrategias de erradicación de la pobreza.
- ✓ Erradicación de la violencia en mujeres y niñas. En este sentido, ONU Mujeres apoya a los gobiernos en la elaboración y aplicación de leyes y estrategias para prevenir todas las formas de violencia contra las mujeres y niñas y responder a ellas, el acceso a servicios de salud sexual y reproductiva para las víctimas y el establecimiento de normas de calidad para la prestación integral de servicios multisectoriales y centrados en la víctima.
- ✓ Empoderamiento de las mujeres, en la paz, la seguridad y la acción humanitaria. Esta área pretende que haya una mayor incorporación de compromisos sobre la igualdad de género en las respuestas durante los conflictos y después de éstos a través de: promocionar que haya más mujeres en las conversaciones de paz, mayor empoderamiento de la mujer en los acuerdos y los compromisos de paz, mejores servicios para atender amenazas a la seguridad contra las mujeres y la investigación y el procesamiento de delitos de género; y, en general, promover todas aquellas iniciativas para fortalecer los medios de vida de las mujeres en actividades de recuperación económica.
- ✓ Fortalecimiento de los sistemas de rendición de cuentas nacionales sobre compromisos relativos a la igualdad entre los géneros. A través de esta área prioritaria, se pretende apoyar al fortalecimiento de la integración de las perspectivas de género en las estrategias de desarrollo y planes sectoriales nacionales; al incremento de los recursos destinados a la igualdad de género y los derechos de las mujeres; a crear una mayor conciencia de sus derechos por parte de las mujeres que viven con el VIH o se ven afectadas por éste.
- ✓ Fortalecimiento de la integración y coherencia de procesos normativos intergubernamentales. ONU Mujeres colabora en el afianzamiento de la normativa mundial y el marco de políticas para la igualdad entre los géneros y el empoderamiento de la mujer y su aplicación; apoya que los procesos sectoriales intergubernamentales reflejen las perspectivas de género; promueve iniciativas que estrechen los vínculos y sinergias entre los marcos normativos y las actividades operacionales en favor de la igualdad entre los géneros y el empoderamiento de la mujer.

## Gobernabilidad

El órgano rector de las actividades operacionales de ONU Mujeres es la Junta Ejecutiva. Está formada por 41 Estados Miembros elegidos por el Consejo Económico y Social de las NU (ECOSOC), por periodos de tres años, y con una distribución en base a 6 grupos regionales. Se reúne anualmente y, entre sus funciones, están la de aprobar el Plan Estratégico, Planes Corporativos, revisión y aprobación de presupuestos, examinar los Informes Anuales y velar por una eficiente priorización de actividades y optimización de recursos<sup>165</sup>.

La Dirección Ejecutiva es el órgano máximo responsable ante la Junta de la ejecución de los dos Subprogramas existentes: 1) Apoyo intergubernamental, coordinación y alianzas estratégicas, y 2) Políticas y Programas. Su estructura también incluye una División de Gestión y Administración, Recursos Humanos y la Oficina de Evaluación. La Sede está en Nueva York. Tiene 6 Oficinas Regionales, 6 Oficinas que cubren varios países, 48 Oficinas a nivel nacional y presencia en cerca de 30 países.

## Recursos

El 98% de los recursos de ONU Mujeres proviene de contribuciones voluntarias, principalmente de los gobiernos, y el 2% del presupuesto regular de NU. A finales de 2014, un total de 143 países han aportado financiación para la realización de las actividades del marco programático de ONU Mujeres. La lista de los 10 principales donantes en 2014 incluye Finlandia, Reino Unido, Noruega, Suiza, Dinamarca, Suecia, Canadá, Australia, EEUU y Holanda.

El Presupuesto de la organización para el Bienio 2014-2015 asciende a \$716 millones, de los cuales un total de 15,7 Millones USD provienen de presupuesto regular<sup>166</sup>. ONU Mujeres también administra y moviliza recursos para el Fondo para la Equidad de Género y Fondo Fiduciario de Naciones Unidas para Eliminar la Violencia la Mujer.

## Función de Evaluación

Un año después del establecimiento de ONU Mujeres en 2011, la Junta Ejecutiva aprobó la Política de Evaluación en 2012, donde se definen los mecanismos y principios de la función de evaluación en la organización. La función de evaluación en ONU Mujeres esta principalmente descentralizada (cerca de un 90% de las evaluaciones se realizan desde las oficinas nacionales o regionales). La Política de Evaluación define los roles, responsabilidades y mecanismos de planificación y seguimiento para cada uno de los tipos de evaluaciones que se realizan, al igual que los procedimientos establecidos para el control de calidad de las evaluaciones.

---

<sup>165</sup>UNW/2011/6

<sup>166</sup>A/68/6

El Plan de Evaluación de ONU Mujeres, que incluye las evaluaciones corporativas y descentralizadas, está alineado con el Plan estratégico de la Organización para 2014-2017. En base a él se realiza el Plan de Evaluación Corporativo.

La Dirección de la Oficina de Evaluación Independiente reporta directamente al Director/a Ejecutivo/a de la organización, y tiene canal directo de información a la Junta Ejecutiva donde presenta el Informe Anual de Evaluación.

La función de evaluación en ONU Mujeres está orientada fundamentalmente hacia cuatro áreas:

- ✓ Realización de evaluaciones corporativas
- ✓ Realización de evaluaciones descentralizadas
- ✓ Promoción de una evaluación responsable en género en el sistema NU
- ✓ Fortalecimiento de las capacidades nacionales de monitoreo y evaluación con equidad de género.

La responsabilidad de la Oficina de Evaluación está especificada en la Política de Evaluación, y abarca las cuatro áreas arriba mencionadas. Incluso en las evaluaciones descentralizadas, la Oficina de Evaluación, asesora, coordina, y proporciona conocimiento técnico, a través de los y las especialistas en evaluación ubicados en las Oficinas Regionales<sup>167</sup>.

### Proceso de revisión por pares

La Política de Evaluación de 2012 especifica el requisito de someter la función de evaluación de ONU Mujeres al proceso de revisión por pares en 2014, con el fin de orientar las recomendaciones para la revisión periódica de la Política de Evaluación en 2015.

El proceso de revisión tuvo lugar entre Marzo y Junio de 2014. Para ello, la metodología utilizada incluyó:

- ✓ Análisis de la información existente sobre la creación y establecimiento de ONU Mujeres, su función de evaluación.
- ✓ Un total de 34 encuestas que incluyeron a la gerencia, personal de las oficinas técnicas y de Oficina de Evaluación, tanto a nivel de sede como de oficinas de campo, y también a consultores/as externos/as que habían formado parte de los equipos de evaluaciones recientes.
- ✓ Un cuestionario en línea a los y las representantes de las oficinas de ONU Mujeres en los países,

---

<sup>167</sup> Fuentes: <http://www.unwomen.org/en/about-us/evaluation> y UNEG/UN Women 2014

- ✓ Una visita a la sede en Nueva York donde el panel de revisión por pares tuvo oportunidad de complementar la información con teleconferencias, entrevistas a miembros de la Junta Ejecutiva, a la Directora Ejecutiva, y e incluso a miembros de UNEG.
- ✓ Análisis de una muestra de nueve informes de evaluación para comprobar su calidad, en base a los criterios y estándares de calidad UNEG<sup>168</sup>.

El borrador del informe fue presentado y discutido con la gerencia ejecutiva y el Director de Evaluación. El informe final es público y accesible a través de la página web de UNEG (UNEG/UN Women 2014):

<http://www.uneval.org/document/detail/1626>

---

<sup>168</sup>Los nueve informes de evaluación son: Contribution of UN Women to Prevent Violence Against Women and Expand Access to Services; Contribution to Increasing Women's Leadership and Participation in Peace and security and Humanitarian Response; Gender-responsive budgeting (GRB) India; Mid Term Evaluation One UN Joint Programme on Gender Equality, Uganda; Evaluation of the UN/Government of Liberia Joint Programme on Gender Equality and Women's Economic Empowerment; UNW Anti Human Trafficking Programme; Establishment, Rehabilitation and Activation of Eight Women's Centres in the Gaza Strip and West Bank; The Gender and Governance Program- Kenya; The Gender and Governance Program- Afghanistan.



## Anexo 2: Descripción de puntuaciones en escala difusa por agencia

Condiciones	Factor clave	UNO Habitat	Referencia
<b>Capacidad de aprendizaje</b>	Diseminación de la información evaluativa	Existe acceso público, pero no se practica consistentemente una estrategia de diseminación de la información. En ocasiones, se realizan talleres y presentaciones de manera ocasional. Limitados reportes de síntesis de evaluaciones o de análisis de lecciones aprendidas	vi, 12, 24, 25
	Participación en la evaluación	El proceso de consulta es consistente en todo el proceso evaluativo, incluyendo ToRs e informes. Las sugerencias son tomadas en cuenta, sin comprometer la independencia. Sin embargo se podría mejorar el proceso de consulta a las oficinas descentralizadas.	vi, 19, 25
	Motivación/cultura evaluativa	Entrevistas confirman que están familiarizados con el propósito de la evaluación y sus potenciales usos. La cultura evaluativa de UNO Habitat se ha fortalecido con la introducción del MTSIP en 2010, a través de la alineación de la programación con RBM. En el actual sistema, existe un área denominada innovación que permite capturar conocimiento y lecciones aprendidas en el ciclo programático. En base a las entrevistas, el aspecto del aprendizaje es más fuerte que la rendición de cuentas.	13, 29
<b>Institucionalización</b>	Independencia	Los resultados de la evaluación no son comunicados directamente por el jefe de MEU a la alta gerencia o a los órganos de gobierno, son reportados por el Director Adjunto. La nueva política de evaluación, aprobada en 2013, separa la unidad de evaluación de monitoreo, y el designa al responsable de evaluación con autoridad para reportar a los órganos de gobierno.	v, 6, 20
	Recursos	La asignación de recursos para evaluación no es transparente ni independiente. Los datos son difícilmente accesibles, y no diferenciados entre monitoreo y evaluación. Se considera que la parte de la evaluación esta marginada con respecto a la parte de monitoreo, a la que se dedica la mayoría del tiempo y los recursos. La unidad no tiene control sobre el presupuesto para evaluación. La implementación del plan de evaluación no está asegurada financieramente.	v, vi, 19
	Personal	Existe balance entre senior y junior staff, personal interno y consultores. Los consultores facilitan el conocimiento técnico. El incremento del personal ha sido importante en los últimos años, pero aún es insuficiente.	13
<b>Políticas y prácticas</b>	Plan de evaluación	Las evaluaciones son planificadas y priorizadas durante la formulación del programa del trabajo de la organización. Incluye consultas con los órganos de gobierno, personal, y donantes. El criterio es multidimensional e incluye la relevancia estratégica y el valor innova TiVo. El proceso consultivo culmina con una lista final aprobada por los órganos de gobierno	vi, 11
	Política de evaluación	Hasta el 2013 no existe política de evaluación. Sin embargo existe manual de evaluación, y guías de monitoreo y evaluación publicadas en 2003. No existe una clara definición de roles y responsabilidades para la planificación, implementación y gestión de evaluaciones. La política de evaluación se formula en el 2013 y está alineada con los estándares UNEG.	ES pág. vi, 11, 16
	Sistema de seguimiento a las recomendaciones	Existe un sistema de seguimiento de las recomendaciones. Las divisiones deben reportar a la unidad de evaluación sobre el estado de implementación de las recomendaciones. No se reporta a los órganos de gobierno. Las recomendaciones son incorporadas en la planificación del nuevo ciclo. El sistema de programación lo permite desde 2010.	vi, 14, 19

Condiciones	Factor clave	UNO Habitat	Referencia
<b>Calidad</b>	Informes y estándares	En general la calidad es alta y en línea con los estándares UNEG. Los pagos solo son procesados una vez se comprueba que la calidad del informe es satisfactoria.	20
	Control de calidad	El proceso de evaluación es riguroso. Existe un sistema de control de calidad, y funciona de manera consistente para todas las evaluaciones.	vi, 14
	Valoración positiva del personal	Existe en general satisfacción acerca de la calidad de las evaluaciones, y en general de la información evaluativa generada por MEU.	23
<b>Uso</b>	Generación de cambios.	Clara percepción de que las evaluaciones tienen influencia, y que los hallazgos y las recomendaciones a menudo sirven en la planificación del nuevo ciclo. Sin embargo, existe poca atención a evaluaciones de tipo estratégico, y por tanto los cambios son a menudo operativos y programáticos, pero no a nivel de políticas o estrategias. En general se considera que la evaluación ha sido influyente en la promoción de cambios a nivel organizativo, y en la validación o ajuste de orientaciones operativas (programas y proyectos).	v, vi
	Implementación de recomendaciones	Alto porcentaje de recomendaciones implementadas	26
	Presencia en la toma de decisiones	Las evaluaciones son discutidas en los foros de alta gerencia y órganos de gobierno. La gerencia incide en la importancia de tomar en cuenta la evaluación en la toma de decisiones.	31

Condiciones	Factor clave	PNUD	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	Acceso público a evaluaciones centralizadas y descentralizadas. Se practica una estrategia de diseminación de la información. Para las evaluaciones de país, o proyecto, la información está disponible en el Evaluation Resource Centre. UNDP network y el Organization Evaluation Forum también son utilizados como canales de diseminación. Se han publicado 'diggests' de informes de evaluación. Los informes a la Junta de Gobierno incluyen análisis de síntesis.	26, 47, 49
	Participación en la evaluación	El grado de consulta en aspectos clave como los ToRs y el criterio de evaluación es inconsistente. La evaluación es principalmente esta descentralizada. En ocasiones no hay claridad de los roles (informante, consultivo) y esto crea confusión. Se podría mejorar el nivel de consulta en los países.	11, 30, 39, 42, 46, 51
	Motivación/cultura evaluativa	Conocimiento y sintonía con los propósitos de la evaluación. A nivel de Junta Ejecutiva existe más confusión entre monitoreo y evaluación, pero se coincide en la necesidad de contar con información evaluativa para la toma de decisiones. A nivel de UNDP interno, el propósito del aprendizaje prima sobre el de la rendición de cuentas	30, 31, 43
<b>Institucionalización</b>	Independencia	EO es independiente funcionalmente. Reporta al Administrador del PNUD, y a la Junta a través del Administrador. No hay canal de reporte directo con el EB. El administrador presenta dos tipos de informes al EB sobre evaluación. Se incluyen síntesis, datos sobre seguimiento, plan de evaluación, y resultados de evaluaciones corporativas. En ocasiones se deja proporciona información sobre evaluaciones de país, pero de manera adhoc.	9, 10, 20
	Recursos	La asignación de recursos proviene de 5 fuentes, y no hay un consolidado de las mismas. La Junta Ejecutiva no recibe el plan de evaluación completo. Las evaluaciones de país y proyecto no tienen siempre el presupuesto garantizado. El unidad no tiene capacidad para realizar todas las responsabilidades de la política de evaluación (control de calidad, seguimiento de las recomendaciones)	13, 23, 24, 33
	Personal	Insuficiente para cumplir con el plan de evaluación. Los consultores facilitan el conocimiento técnico, pero en ocasiones se observan importantes deficiencias en la expertise de los equipos.	34,35
<b>Políticas y prácticas</b>	Plan de evaluación	No hay un criterio de aplicación consistente en la selección de áreas para evaluar. La Junta no aprueba el plan de evaluación. Aspectos a mejorar una vez aprobada la política de 2006	22
	Política de evaluación	Borrador que será aprobado en Junio de 2006 por la Junta. Existe un manual desde el 2002-handbook for monitoring and evaluation for results-,	19, 29
	Sistema de seguimiento a las recomendaciones	Sistema de seguimiento incipiente, solo para evaluaciones (corporativas). Dicho sistema presenta muchas deficiencias (no se identifican responsables, periodo de implementación). Por tanto la información sobre recomendaciones que se lleva a la Junta es muy limitada, y no se usa para la formulación del nuevo ciclo. La unidad de evaluación cuestiona que tenga suficiente capacidad para poder llevar a cabo labores de seguimiento de todas las evaluaciones.	19, 27, 47

Condiciones	Factor clave	PNUD	Referencia
<b>Calidad</b>	Informes y estándares	El panel identifica limitaciones a la calidad de los informes. Estas incluyen no link entre hallazgos, conclusiones y recomendaciones, falta de argumentación de los criterios evaluativos a utilizar, y abundancia de recomendaciones no críticas. La divergencia de opiniones no queda reflejada en los informes.	30, 31, 40
	Control de calidad	El control de calidad se considera deficiente y es uno de los elementos clave de la política de evaluación, una vez se apruebe. También se duda de la capacidad de la unidad de proveer asistencia de manera global a todas las oficinas de campo. El apoyo de las oficinas regionales se considera clave, y todavía insuficiente.	13, 31, 33, 40, 51
	Valoración positiva del personal	En general existe una valoración positiva de la calidad de los informes, incluida la Junta. El interés de la Junta en la evaluación ha aumentado en los últimos años.	45, 49
<b>Uso</b>	Generación de cambios.	El panel ha identificado casos de cambios estratégicos debido a información evaluativa. La creciente expectación de la Junta por la evaluación favorece positivamente su relevancia en la toma de decisiones.	19, 45, 52
	Implementación de recomendaciones	El sistema es incipiente y todavía tiene muchas limitaciones. Se considera que el nivel de empoderamiento del personal en la implementación de recomendaciones es muy limitado. Sin embargo, esto puede cambiar con la aprobación de la política de evaluación, puesto que la Junta ha solicitado la formalización de un sistema para el seguimiento a las recomendaciones	19, 27, 47
	Presencia en la toma de decisiones	Las evaluaciones, en particular las corporativas, son discutidas en foros de alta gerencia y Junta. Sus resultados son relevantes en la toma de decisiones.	11, 44, 45

Condiciones	Factor clave	OIOS	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	No se practica consistentemente una política de diseminación de la información. Puesto que no todos los informes de evaluación son presentados a los órganos gubernamentales, no todos son accesibles en la web. Y cada vez el número de informes es menor, por lo que la accesibilidad ha disminuido.	56, 60, 61
	Participación en la evaluación	Participación inconsistente, y en general limitada a solo algunas partes de la evaluación, en particular en los TdRs.	46
	Motivación/cultura evaluativa	Existe confusión entre los representantes del Quinto Comité de la Asamblea General sobre el valor y las diferencias entre evaluaciones con otras funciones. No existe un claro entendimiento sobre lo que la evaluación puede aportar. La evaluación es confundida con auditoría, inspecciones y el general con control y rendición de cuentas.	30, 39, 60, 61
<b>Institucionalización</b>	Independencia	OIOS trabaja con independencia plena a la hora de reportar a los órganos de gobierno. Reporta a la Asamblea General a través del SG. El mandato asegura imparcialidad e independencia. La cultura de independencia evaluativa es muy fuerte en la organización. Esto puede haber llevado a una situación de aislamiento a la unidad de evaluación.	31, 39
	Recursos	El presupuesto ha sido incrementado. La independencia presupuestaria es muy alta. Los datos son transparentes y disponibles. Sin embargo, el panel considera que puede llegar a ser insuficiente si el mandato de OIOS requiere cubrir todas las actividades de la Secretaría.	30, 33
	Personal	Cerca del 95% de los puestos provienen del presupuesto regular. El incremento del personal ha sido importante en los últimos años. Balance entre senior y junior staff. Sin embargo, se tiende a prescindir de consultores (restricciones presupuestarias) y en ocasiones el equipo no cubre todas las necesidades técnicas	43, 44
<b>Políticas y prácticas</b>	Plan de evaluación	El plan es preparado por la unidad en base a un criterio de riesgo. Este criterio asegura transparencia en la selección de áreas, y relevancia y puntualidad de las evaluaciones. El proceso de consulta existe, aunque es algo limitado y rígido. Las necesidades del personal son tomadas en cuenta, al igual que las demandas de los Estados Miembros. Es aprobado por el Quinto Comité de la Asamblea General.	27, 32
	Política de evaluación	La PPBME, no es exactamente una política al uso, pero refleja el mandato y los procedimientos a seguir. En principio se adhiere a las normas UNEG, pero puesto que su formulación fue anterior a ella, necesitaría de una actualización, en particular para reflejar mejor las diferencias entre evaluación, inspección y auditoría.	41
	Sistema de seguimiento a las recomendaciones	Existe un sistema de dar seguimiento a las recomendaciones. El sistema Issue Track permite la revisión on line del estado de las recomendaciones. Se actualiza cada 6 meses. Además, la revisión trienal facilita el seguimiento de proyectos acabados a los órganos de gobierno. Sin embargo, el sistema no permite una revisión cualitativa de las recomendaciones ni una consideración completa de todas las evaluaciones por los órganos de gobierno (puesto que no todas son reportadas a la AG). El sistema está muy orientado a la rendición de cuentas	37, 56

Condiciones	Factor clave	OIOS	Referencia
<b>Calidad</b>	Informes y estándares	En general la calidad es alta, a pesar de que en ocasiones no cumplan todos los estándares de calidad UNEG. Tienden a ser informes demasiado cortos, y es difícil percibir cual es la evidencia que justifica los hallazgos.	47-49, 54
	Control de calidad	Existe un proceso documentado de control de calidad que da seguimiento a todas las fases de la evaluación. Este proceso de control de calidad ha contribuido a mejorar la calidad de los informes de evaluación.	46
	Valoración positiva del personal	Satisfacción sobre la calidad de los informes de evaluación. Percepción interna de que la calidad ha mejorado.	47
<b>Uso</b>	Generación de cambios.	Limitada evidencia de que la evaluación haya promovido cambios, o haya servido para la mejora de estrategias y programas. Sin embargo la tendencia ha mejorado, y se han reportado algunos casos (A/63/221, y OCHA)	58, 59
	Implementación de recomendaciones	No existe un sistema de seguimiento completo que permite analizar el porcentaje de recomendaciones implementadas. Sistema no orientado al aprendizaje sino a la rendición de cuentas.	37, 56, 59
	Presencia en la toma de decisiones	Los procesos y foros donde las evaluaciones son analizadas no son los más adecuados para la toma de decisiones. El Quinto Comité de la Asamblea General discute principalmente temas de presupuesto. No existen oportunidades claras de usar la evaluación en los foros de decisión.	29, 59

Condiciones	Factor clave	PNUMA	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	Existe acceso público a las evaluaciones. Aunque no exista una política formal de diseminación, en la práctica los hallazgos de las evaluaciones son regularmente presentados a través de talleres. Además, la unidad mantiene una base de datos de lecciones aprendidas derivadas de las evaluaciones. Se realizan reportes de síntesis de evaluaciones bianuales. Los reportes bianuales son presentados en órgano ejecutivo, y diseminados a gobiernos, personal y Consejo.	20, 25, 31
	Participación en la evaluación	Buen nivel de consulta y participación en las evaluaciones. Sin embargo existe la percepción de que no siempre se involucra a los actores clave en los países, particularmente en la planificación y discusión de los hallazgos	19, 23, 25, 30
	Motivación/cultura evaluativa	La visibilidad y relevancia de la evaluación ha aumentado en los últimos años. Se considera que la alta gerencia promueve la cultura evaluativa, y la incluye en la agenda de la toma de decisiones	Vi, 16, 20
<b>Institucionalización</b>	Independencia	A pesar de estar contemplado en la política de evaluación, no existe un canal de comunicación directa a los órganos de gobierno. Existe una independencia estructural de la función de evaluación.	Vi vii, 20
	Recursos	El presupuesto regular cubre salarios de persona, y gastos administrativos y viajes. 50% bajo el control de las unidades técnicas. La mayoría del presupuesto está asignado a las evaluaciones GEF.	V, vi, 17, 18
	Personal	Existe balance (algo insuficiente) entre personal y consultores, estos últimos hacen gran parte del trabajo técnico. Por mejorar el balance entre consultores nacionales e internacionales.	17, 24
<b>Políticas y prácticas</b>	Plan de evaluación	Plan bianual en base a un criterio establecido de fase de terminación y presupuesto. La mayoría de las evaluaciones son las mandatorias GEF. Existe transparencia y predictibilidad de las evaluaciones. El nivel de consulta se considera adecuado. Aprobado a nivel interno (ED)	19, 30
	Política de evaluación	EP formalmente aprobada en 2009 por el Director Ejecutivo, en base a normas UNEG. Fue presentada al CPR pero nunca fue sometida formalmente al GC para su aprobación. Algunas premisas de la política no se cumplen, como por ejemplo el canal de reporte directo.	16
	Sistema de seguimiento a las recomendaciones	Existe un sistema de seguimiento a las recomendaciones. Los managers reportan sobre el estado de las recomendaciones cada 6 meses. La unidad analiza la información y se discute a nivel de alta gerencia. No se reporta a los órganos de gobierno. Por mejorar los niveles de empoderamiento del personal en la implementación.	Vii, 32, 33

Condiciones	Factor clave	PNUMA	Referencia
<b>Calidad</b>	Informes y estándares	La calidad es alta, en línea con los criterios de calidad UNEG. Las metodologías son sólidas y a menudo innovadoras. Los pagos solo son procesados una vez se comprueba que la calidad del informe es satisfactoria.	Vi, 23, 24, 25, 26
	Control de calidad	Existe un sistema de control del calidad, y salvo algunas excepciones, se considera que funciona de manera consistente para todas las evaluaciones	Vi, 23
	Valoración positiva del personal	En general, existe satisfacción por la calidad de los informes y en general la información generada por la evaluación.	23, 27
<b>Uso</b>	Generación de cambios.	Clara percepción de que las evaluaciones son útiles en la formulación de nuevos proyectos. No tantos cambios a nivel estratégico, o de políticas, pero si operativos. A nivel estratégico, la evaluación del Subprograma de la División de Tecnología, Industria y economía, y la Evaluación formativa del Programa de Trabajo 2010-2011 se considera que han sido útiles para la toma de decisiones a nivel de programas y políticas.	Vi, 30
	Implementación de recomendaciones	Alto porcentaje de recomendaciones implementadas, cerca del 70%	32
	Presencia en la toma de decisiones	Gerencia apoya de manera directa la evaluación, la promueve. Se usan las estadísticas del estado de las recomendaciones en la valoración del rendimiento de los directores de los programas. Existen pocas oportunidades de presentar la evaluación a los órganos de gobierno, y esto limita su potencial	Vi, 16, 30, 31

Condiciones	Factor clave	UNICEF	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	Existen algunos procedimientos de comunicación como acceso público a los informes no se practica una estrategia de diseminación de la información de los informes.	9
	Participación en la evaluación	Participación limitada de los actores clave en la evaluación, más allá de propósitos informativos.	8
	Motivación/cultura evaluativa	El liderazgo de UNICEF en evaluación es muy alto y tiene alta credibilidad; especialmente se valora su contribución en armonización de normas y buenas prácticas. En general existe un entorno favorable que facilita el aprendizaje, pero cohabita con un gran desnivel entre HQs y oficinas de campo.	8, 9
<b>Institucionalización</b>	Independencia	Hasta la aprobación de la política de evaluación en 2007 existía confusión en las líneas de reporte de la función de evaluación. Se asumía que el jefe de evaluación tenía cierta autoridad, pero no estaba delimitado explícitamente	7
	Recursos	Limitada predictibilidad presupuestaria. Al menos un tercio no forma parte del presupuesto regular y debe ser negociado con los programas. No hay transparencia ni datos claros de porcentajes destinados a evaluación en las oficinas de campo	5
	Personal	Gran desnivel entre HQs y oficinas de campo. HQs cubre el 3% de la evaluación existente. En las oficinas de campo, se considera que la mitad personal con funciones de M&E. Los conocimientos de M&E en las oficinas de campo son limitados.	8
<b>Políticas y prácticas</b>	Plan de evaluación	Con anterioridad a la PE, no existía un plan de evaluación, ni un criterio específico de selección de las evaluaciones. Proceso muy liderado por donantes. Después de la aprobación de la política, se ponen en marcha criterios transparentes rigurosos y vinculados al plan estratégico	5, 8
	Política de evaluación	La política de evaluación fue aprobada por la Junta Ejecutiva en 2008, y revisada en 2013.	6
	Sistema de seguimiento a las recomendaciones	Incipiente por en la sede, aun por introducir en las oficinas de campo. No es utilizado para reportar sobre el estado de las recomendaciones a los órganos de gobierno.	10

Condiciones	Factor clave	UNICEF	Referencia
<b>Calidad</b>	Informes y estándares	Informe reciente mencionado en el PR identifica pobres resultados en calidad. Importantes esfuerzos realizados desde la sede por armonización.	9
	Control de calidad	Control de calidad limitado y desigual, no estandarizado. Capacidades muy limitadas a nivel de oficinas de campo para poder realizar controles de calidad. Gran mejora en la revisión de 2013 gracias al sistema GEROS	8
	Valoración positiva del personal	Personal con conocimiento limitado de lo que se genera. Existe la percepción entre los Estados Miembros que la información reportada a los órganos de gobierno no es sustantiva ni analítica.	10
<b>Uso</b>	Generación de cambios.	Existen indicios y algunas indicaciones de cambios potenciales. La limitada recogida de información sobre recomendaciones y en general sobre información evaluativa limita el análisis de evidencias, más allá de indicaciones. (pág. 9, 10)	9, 10
	Implementación de recomendaciones	No existe un sistema de seguimiento de las evaluaciones de las oficinas de campo (97%) del total. Con respecto a HQs el sistema es incipiente, aunque no se usa para reportar a los órganos de gobierno. El sistema ha mejorado, aunque todavía se observan deficiencias en su aplicación (pág. 37)	10
	Presencia en la toma de decisiones	La presencia y sustantividad de la información evaluativa a los órganos de gobierno es limitada, tal y como el panel ha apuntado. Existe poca oportunidad de utilizar la evaluación en la toma de decisiones debido a su escasa diseminación, y limitada presencia en foros de decisión. Esto mejora en la revisión de 2013, parcialmente por GEROS (pág. 48)	10

Condiciones	Factor clave	FAO	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	La diseminación de las evaluaciones de país y de proyecto es limitada a la publicación del informe en la web. El panel incide en la necesidad de mejorar los mecanismos de retroalimentación con las unidades operativas, de cara a contribuir al aprendizaje organizativo.	39
	Participación en la evaluación	El proceso de consulta inconsistente, en particular en las evaluaciones de país y proyecto. Se han establecido en ocasiones grupos consultivos.	Xi, 23, 46
	Motivación/cultura evaluativa	La cultura evaluativa es sólida. La unidad de evaluación lleva establecida desde 1968. Tanto los managers como los miembros del Consejo coinciden en la valoración de la evaluación como instrumento de mejora.	29, 42
<b>Institucionalización</b>	Independencia	Unidad independiente, no relacionada con las unidades operacionales. Tiene canal directo de reporte al Comité de Programas, que es un sub-comité del Consejo. El Comité de Programas aprueba el plan de evaluaciones a los órganos de gobierno, recibe en informe de dichas evaluaciones y revisa el estado de las recomendaciones.	11, 13, 49
	Recursos	Existe criterio transparente de asignación a evaluación. En general la oficina tiene datos disponibles y transparentes para predecir cuál va a ser su presupuesto y planificar sus actividades, bien sea presupuesto regular o extra presupuestario.	14-16
	Personal	Mayor temporalidad debida al aumento de asignación extra presupuestaria Consultores y expertos son contratados para hacerse cargo del conocimiento técnico de las evaluaciones.	17-19
<b>Políticas y prácticas</b>	Plan de evaluación	La planificación sigue procesos diferentes, en función del tipo de evaluación. En algunos casos, Comité de Programas hace la selección. Para las evaluaciones de país, la EO ha establecido un criterio multidimensional que incluye la relevancia del factor de la agricultura. Esta pre-selección es endosada por el Comité de Programas. Para las evaluaciones de proyectos, EO propone a las unidades programáticas.	21-22
	Política de evaluación	No existe documento de política en sí. (Charter). Formalmente aprobado por el Consejo, sigue estándares UNEG. Su elaboración es consecuencia del Plan de Acción Inmediata aprobado en 2008	X
	Sistema de seguimiento a las recomendaciones	Existe un sistema de seguimiento, aplicación inconsistente. Los órganos de gobierno reciben informes de ciertas evaluaciones dirigidas a ellos. No se analiza el estado de las recomendaciones de las evaluaciones de país o de proyecto.	24, 25

Condiciones	Factor clave	FAO	Referencia
<b>Calidad</b>	Informes y estándares	La calidad de los informes inconsistente. A menudo la metodología aplicada, las limitaciones, y en general los criterios evaluativos no están explicados.	Xi, 44, 61
	Control de calidad	El control de calidad no se aplica consistentemente en las evaluaciones. La EO no asume responsabilidad por los contenidos, ni por la calidad de sus informes.	Xi, 23
	Valoración positiva del personal	La credibilidad de los informes depende en gran medida de la credibilidad de los managers de evaluación a nivel personal e individual.	41-42
<b>Uso</b>	Generación de cambios.	Percepción general que la evaluación no se usa en la toma de decisiones ni generación de cambios. Los miembros del PC consideran que la información no es analítica y dificulta el link entre evaluación y programación.	29, 31
	Implementación de recomendaciones	No existe un sistema de seguimiento de las evaluaciones de país, ni de proyecto. Las recomendaciones no encajan en el ciclo programático, y hay confusión en las responsabilidades en la implementación de recomendaciones.	24, 25, 35
	Presencia en la toma de decisiones	Escasa influencia. Información no ligada a aspectos estratégicos. No existe un foro de discusión de evaluaciones de país y proyecto.	X, 32, 34-35

Condiciones	Factor clave	ONUDI	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	Existe acceso público, pero no existe una estrategia de diseminar la información evaluativa de manera consistente. En ocasiones se hacen presentaciones, pero la participación es limitada e irregular (extractos de lecciones aprendidas al PAC, u otros comités de personal).	35, 50
	Participación en la evaluación	En general tanto los ToRs como los informes son consultados con los actores clave. Se podría mejorar los niveles de cobertura, especialmente las oficinas de campo.	44
	Motivación/cultura evaluativa	Confusión entre monitoreo y evaluación, y en general dudas sobre el valor añadido de la evaluación en la programación. Cierta nivel de frustración sobre los resultados esperados en la planificación de los proyectos, y en el rol de la evaluación en buscar impacto. El cuestionario revela que solo una minoría de los managers da importancia a la evaluación.	9, 31, 49, 50
<b>Institucionalización</b>	Independencia	EVA es estructuralmente independiente. Reporta al DG y esta desvinculada de las unidades temáticas. Tiene canal directo con el IDB (órgano de gobierno) de manera bianual para presentar actividades de evaluación.	8, 21, 27
	Recursos	Presupuesto insuficiente. Las evaluaciones de los proyectos, corren totalmente a cargo de presupuesto de los proyectos, fuera del control de EVA.	8, 22-23, 33
	Personal	Se recurre en exceso a la contratación de consultores externos para tareas sustantivas. Esto incide negativamente en la credibilidad del trabajo de EVA, ya que los consultores no son considerados tan imparciales.	44
<b>Políticas y prácticas</b>	Plan de evaluación	Existen procedimientos transparentes y consultivos en la preparación del plan de trabajo. Finalmente es aprobado por el EB (interno). El criterio de selección incluye presupuesto, extensiones, requerimientos de donantes, anticipación de problemas, y áreas innovadoras. Podría mejorar en nivel de consulta con las oficinas de campo.	23, 25, 33, 40
	Política de evaluación	La política de evaluación fue aprobada por el Director General y el Comité Ejecutivo en el 2006. En general, en línea con UNEG. Muy poca relevancia entre el personal (la mayoría no la conoce), y presenta ciertas inconsistencias con otras guías como las de cooperación técnica.	9, 23, 38
	Sistema de seguimiento a las recomendaciones	Existe un sistema de seguimiento. Se reporta sobre su cumplimiento al EB (comité interno). Su uso es muy cuestionado por los managers, poca credibilidad. Limitada evidencia de su uso para implementación o formulación de nuevos proyectos.	10, 26, 47-48

Condiciones	Factor clave	ONU DI	Referencia
<b>Calidad</b>	Informes y estándares	La calidad de los informes inconsistente. La información sobre la metodología, y criterios aplicados no es explicada. Poca importancia otorgada al criterio de relevancia. Inconsistente presentación de datos que justifique los hallazgos. Esto también es debido a la debilidad de la planificación y monitoreo de los proyectos.	9, 43
	Control de calidad	No existen controles de calidad consistentes. Según el panel, esto explica la variedad en los niveles de calidad de los informes.	9, 49
	Valoración positiva del personal	Opinión diversa. En ocasiones, se alude a que la demasiada presencia de consultores, limita el entendimiento del programa, y la credibilidad del informe (28, 41, 42)	28, 41-42
<b>Uso</b>	Generación de cambios.	Existen cambios operativos, aunque limitada evidencia a nivel estratégico. Casos reportados de cambios incluyen India CSF, Cleaner Production y Evaluation of IP.	50
	Implementación de recomendaciones	Cerca del 50% se implementa, pero los managers no se han empoderado del proceso, del que muchos cuestionan su utilidad.	47,48
	Presencia en la toma de decisiones	Limitada presencia de la unidad en los procesos y foros donde se toman las decisiones a nivel estratégico. El informe bianual al IDB es sobre actividades (no resultados o lecciones aprendidas)	34

Condiciones	Factor clave	FMAM	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	Existe un repositorio electrónico de las evaluaciones públicamente accesible y ciertos productos de diseminación (knowledge tree series). La diseminación es limitada.	16, 41, 47, 72
	Participación en la evaluación	Limitados procesos de consulta. Las evaluaciones son consideradas top down. El tiempo para reaccionar a los informes es considerado muy insuficiente. En general poca participación.	10, 12, 15, 41, 59, 69
	Motivación/cultura evaluativa	Débil empoderamiento de la organización con los resultados de la función de evaluación. Cierta nivel de competitividad entre la Secretaría y la unidad de evaluación. A nivel de país, existe gran desconocimiento de la labor de la unidad de evaluación.	13, 67-69
<b>Institucionalización</b>	Independencia	Estructuralmente es totalmente independiente, reporta directamente al Council, quien elige al Director de la EO. Falta oficializar legalmente esta independencia, pero su aplicación es total. La Secretaría y agencias implementadoras coinciden en afirmar que esta independencia ha aislado la unidad de evaluación, y ha afectado negativamente los procesos de comunicación y consulta.	11, 27-31, 45, 46
	Recursos	La independencia financiera plena. Su presupuesto garantizado para cumplir el plan cuatrienal, el cual se desglosa en planes operativos. En 4 años, el presupuesto ha aumentado cerca del 50%. Suficiente para cumplir con las actividades previstas.	10, 35-37, 48
	Personal	Equilibrado. Se limita la presencia de consultores externos a expertise técnica muy específica con el fin de salvaguardar y armonizar los estándares de calidad. (pág. 39)	39
<b>Políticas y prácticas</b>	Plan de evaluación	El plan de evaluación es preparado por la unidad, en base a un criterio estándar multidimensional establecido, y aprobado por el Council. El proceso de consulta y participación podría mejorar.	15, 47
	Política de evaluación	Política de evaluación aprobada por el Council en el 2006, y actualizada en 2010. En general en línea con UNEG. Como complemento a la EP, la EO ha publicado numerosas guías que dan operatividad a la política evaluativa.	11
	Sistema de seguimiento a las recomendaciones	Existe un sistema de seguimiento. Es considerado como algo rígido por el panel, solo se dan 2 semanas a los managers para presentar el plan de acción. El Management Action Record (MAR) contiene un análisis estadístico de la adopción de decisiones por parte del Council y su relación con las recomendaciones de las evaluaciones. Podría mejorarse la parte cualitativa	12, 58, 66

Condiciones	Factor clave	FMAM	Referencia
<b>Calidad</b>	Informes y estándares	Los informes incluyen sólida información técnica, hallazgos basados en evidencias rigurosas, y análisis profundos. Rigor metodológico en las evaluaciones realizadas. En algunas evaluaciones (evaluaciones de país), el trabajo de campo puede resultar limitado y afectar la calidad de los informes	12, 41, 57
	Control de calidad	Considerados algo suaves por el panel, sobre todo teniendo en cuenta la relevancia del factor técnico de las evaluaciones	14
	Valoración positiva del personal	Valoración positiva de los informes, especialmente por el Council. Su credibilidad viene sobretodo fundamentada por su calidad metodológica.	11, 46
<b>Uso</b>	Generación de cambios.	Cambios reportados, pero el limitado proceso consultivo y escasa participación de la Secretaría y contrapartes hace que su repercusión disminuya. La percepción top down de las evaluaciones no ayuda en la generación de cambios fruto de los resultados de las evaluaciones.	13
	Implementación de recomendaciones	El MAR es un sistema totalmente enfocado al Council. Seguimiento estadístico, difícilmente aplicable para valorar como las recomendaciones afectan la formulación del nuevo ciclo.	66
	Presencia en la toma de decisiones	Miembros del Council positivos con el trabajo de evaluación. El panel ha encontrado ejemplos concretos donde las evaluaciones han sido tomadas en cuenta en foros de toma de decisiones. Existe una percepción positiva de la evaluación como herramienta de información para la toma de decisiones entre los miembros del Council.	47, 65-66, 68

Condiciones	Factor clave	PMA	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	Existe acceso público, pero no se practica una estrategia de diseminación de la información. Al EB solo se presentan sumarios de las evaluaciones corporativas. Las descentralizadas son incluidas en el informe anual, accesible en la Web.	15, 32, 64, 65
	Participación en la evaluación	Participación limitada e inconsistente de los actores clave en la evaluación, principalmente para aspectos informativos	12, 13, 37, 51
	Motivación/cultura evaluativa	La evaluación tiene un perfil muy poco relevante en la organización, y se cuestiona que aporte significativamente valor al rendimiento organizacional. Los conocimientos y la cultura evaluativa en general es débil, particularmente en las oficinas regionales, que son encargados de apoyar tareas evaluativas de las oficinas de campo. Existe confusión entre auto-evaluación y evaluación descentralizada.	37, 42, 60
<b>Institucionalización</b>	Independencia	La función de evaluación existe desde 1965. Desde 2006 el Director de EO reporta directamente al DG, y presenta los informes directamente a los órganos de gobierno	12, 35
	Recursos	La asignación de recursos no es transparente. Evaluaciones no son realizadas por falta de presupuesto.	13, 40
	Personal	Falta de recursos en las oficinas de campo, y limitados conocimientos técnicos (gran desnivel entre sede y oficinas de campo)	51, 52
<b>Políticas y prácticas</b>	Plan de evaluación	El Plan de evaluación lo hace la EO en consulta con las oficinas regionales, es aprobado internamente. Por mejorar aun los criterios utilizados.	32, 36, 40
	Política de evaluación	Falta de política de evaluación, en su lugar, una serie de documentos que se complementan, en ocasiones inconsistentes entre ellos, y donde no se pueden considerar claramente que constituyan una política de evaluación en base a estándares UNEG.	46
	Sistema de seguimiento a las recomendaciones	Incipiente, los managers son responsables de realizar un plan de acción, pero la unidad de evaluación no tiene responsabilidad directa en el seguimiento. Se espera que en el 2007 se introduzca un nuevo sistema.	13, 14, 31, 64

Condiciones	Factor clave	PMA	Referencia
<b>Calidad</b>	Informes y estándares	Calidad inconsistente, especialmente la evaluación descentralizada. Por mejorar el detalle de la metodología,, la claridad en la información y la calidad de las recomendaciones	12, 48, 53, 56, 57
	Control de calidad	No existen controles de calidad consistentes. El apoyo a las oficinas regionales y de campo es ad-hoc. Se considera que los informes de evaluación en ocasiones descansan en información no validada.	30, 50, 51, 53
	Valoración positiva del personal	Percepción negativa de la credibilidad y calidad de los informes, poco rigor metodológico incluyendo la poca aplicabilidad de las recomendaciones.	37, 53
<b>Uso</b>	Generación de cambios.	Limitadas indicaciones de que las evaluaciones tengan impacto y/o generen cambios aunque sea a nivel operativo.	13, 62
	Implementación de recomendaciones	Limitado empoderamiento por parte de los managers por la implementación de las recomendaciones. El EB tampoco ha otorgado mucha relevancia a la inclusión del reporte de las recomendaciones en los informes que se les presentan.	14, 31
	Presencia en la toma de decisiones	La presencia de las evaluaciones en los foros de decisión es prácticamente inexistente. El PCR es el foro de discusión interno, pero de muy poca efectividad real. Las evaluaciones se centran en aspectos operativos, no estratégicos, y no encajan muy bien en foros de decisión referentes a estrategias y políticas.	13, 37, 42

Condiciones	Factor clave	ONU Mujeres	Referencia
<b>Capacidad de Aprendizaje</b>	Diseminación de la información evaluativa	Existe acceso público y actividad intensa de diseminación de la información evaluativa. Puesto que cerca del 90% de la evaluación es descentralizada, HQs apuesta por la diseminación a través de la web, y el uso de plataformas on line (UNW GATE, GERAAS). Síntesis de evaluaciones son incluidas como parte del informe anual y el primer meta análisis de evaluaciones fue publicado en 2014. El acceso estadístico al estado de las recomendaciones es público y parte de la estrategia de comunicación. Se podría mejorar la diseminación de la información desde las oficinas de campo.	4, 22, 23
	Participación en la evaluación	Se considera que el nivel de participación y consulta de los actores clave en todo el proceso evaluación es satisfactorio, y tiende a involucrar tanto a gerencia, órganos de gobierno y personal interno.	3, 12
	Motivación/cultura evaluativa	Existe un alto nivel de apoyo a la función de evaluación tanto por la Junta, alta gerencia y personal. La unidad está muy involucrada en la armonización de estándares de calidad y en general en actividades de UNEG, y EvalPartners	3, 19
<b>Institucionalización</b>	Independencia	El director de EO reporta directamente al DG, y su canal de reporte con la Junta es directo. El nivel de autonomía e independencia es considerado alto.	16
	Recursos	El presupuesto propuesto por la unidad es aprobado sin modificaciones, y es considerado apropiado y suficiente.	12, 18
	Personal	Cerca del 90% de la evaluación es descentralizada, y en gran parte la evaluación es realizada por los managers/consultores. Existe riesgo de que esto limite la credibilidad de las evaluaciones.	4, 21
<b>Políticas y prácticas</b>	Plan de evaluación	El plan de evaluación sigue criterios multidimensionales establecidos en la política de evaluación tanto para evaluaciones corporativas y descentralizadas, y goza de un alto nivel de consulta, participación y consenso. El plan estratégico se establece para 3 años, y se traduce operativamente en el plan corporativo. El plan forma parte del informe que se presenta a la Junta para su aprobación.	7, 20
	Política de evaluación	Política de evaluación fue aprobada por la Junta en 2012, en línea con UNEG	7, 17
	Sistema de seguimiento a las recomendaciones	Existe un sistema de seguimiento, público y accesible, GATE, al igual que estadísticas sobre el estado de implementación de las recomendaciones.	23

Condiciones	Factor clave	ONU Mujeres	Referencia
<b>Calidad</b>	Informes y estándares	Los informes presentados en el análisis presentan calidad en base a los estándares UNEG, aunque se observan ciertas diferencias entre unos y otros, aunque no grandes deficiencias metodológicas.	20
	Control de calidad	Rigurosos controles de calidad, que incluyen las evaluaciones tipo descentralizado. De todas maneras, el panel considera que se podrían mejorar los controles, teniendo en cuenta la relevancia de la evaluación descentralizada en UNW, y el desbalance de personal experto en M&E en algunas de las oficinas de campo. Existe un comité asesor de evaluación que analiza y vela por la calidad de los informes, y por el establecimiento de controles de calidad rigurosos.	4, 14, 20
	Valoración positiva del personal	Percepción muy positiva de la alta gerencia, personal y junta consultados sobre la calidad de los informes de evaluación, y en general el trabajo desempeñado por la unidad	19
<b>Uso</b>	Generación de cambios.	A pesar de que la valoración es positiva, muy pocos entrevistados pudieron identificar ejemplos de cómo se han utilizado los hallazgos de una evaluación. Se considera que es un proceso que va a llevar tiempo, y que se está en la dirección correcta.	22
	Implementación de recomendaciones	El estado de las recomendaciones es accesible a través de GATE. Actualmente existen cerca de un 70% de recomendaciones implementadas.	23
	Presencia en la toma de decisiones	La evaluación forma parte de la agenda donde se toman las decisiones, las reuniones del EB, y durante la presentación de los informes anuales.	12