



**UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO**

---

---

---

**CENTRO UNIVERSITARIO UAEM TEXCOCO**

**“META-EVALUACIÓN DE LA EVALUACIÓN DE DISEÑO DEL PROGRAMA  
DE INCLUSIÓN SOCIAL PROSPERA 2016”**

## **TESIS**

**PARA OBTENER EL TÍTULO DE LICENCIADA(O) EN CIENCIAS POLÍTICAS Y  
ADMINISTRACIÓN PÚBLICA**

**PRESENTA**

**GONZALEZ CISNEROS DIANA LAURA  
VILLARREAL CORTES JUAN DANIEL**

**DIRECTOR**

**M. EN P.P. ANTONIO INOUE CERVANTES**

**REVISORES**

**DR. ALEJANDRO A. P. AGUILAR MIRANDA  
M. EN G. y A.P. KARINA REBECA RAMÍREZ GONZÁLEZ**

**TEXCOCO, ESTADO DE MÉXICO, 03 DE JULIO DE 2019.**

## ÍNDICE

<b>INTRODUCCIÓN .....</b>	<b>1</b>
<b>CAPÍTULO 1. CONSIDERACIONES TEORICAS DE LA METAEVALUACIÓN.....</b>	<b>4</b>
1.1. Antecedentes teóricos de la Metaevaluación.....	4
1.2 Conceptos y objetivos de Metaevaluación.....	12
1.3 Instrumentos de la Metaevaluación.....	17
1.4 Premisas de la Metaevaluación.....	24
1.5 Tipos de Metaevaluación.....	28
<b>CAPÍTULO 2. METAEVALUACIÓN EN EL MUNDO.....</b>	<b>34</b>
<b>2.1 Experiencias de Metaevaluación en Gran Bretaña .....</b>	<b>36</b>
2.1.1 La Metaevaluación del Sistema de Evaluaciones de EuropeAid .....	36
2.1.2 Metaevaluación de la Agenda de Modernización de los Gobiernos Locales del Reino .....	40
<b>2.2 Experiencias de Metaevaluación en España.....</b>	<b>42</b>
2.2.1 Metaevaluación de políticas de igualdad de género en España.....	43
2.2.2 La evaluación de Programas de Reforma Educativa en España: Tres estudios de caso desde un enfoque de Metaevaluación.....	48
<b>2.3 Experiencias de Metaevaluación en América Latina .....</b>	<b>50</b>
2.3.1 Experiencias de Metaevaluación en Costa Rica .....	50
<b>2.4 Experiencias de Metaevaluación en México.....</b>	<b>52</b>
2.4.1 Metaevaluación del fondo nacional para el fomento de las artesanías 2012.....	54
2.4.2 Metaevaluación del programa de abasto rural a cargo de DICONSA 2007-2012...	56
2.4.3 EVALUA DF 2008.....	58
<b>CAPÍTULO 3. ANÁLISIS DE LA EVALUACIÓN DE DISEÑO DEL PROGRAMA DE INCLUSIÓN SOCIAL PROSPERA 2016.....</b>	<b>64</b>
<b>3.1 Programa de Inclusión Social PROSPERA.....</b>	<b>64</b>
3.1.1 Descripción General del Programa.....	64
<b>3.2 Evaluación de Diseño.....</b>	<b>66</b>
<b>3.3 Características de la Evaluación del diseño del Programa de Inclusión Social “PROSPERA 2016” .....</b>	<b>69</b>

3.3.1 Objetivo.....	71
3.3.2 Criterios.....	72
3.3.3 Metodología.....	74
3.3.4 Contenido / desarrollo.....	76
3.3.5 Conclusiones Generales.....	77
<b>3.4 Conclusión del Análisis de la evaluación de diseño del Programa de Inclusión Social PROSPERA 2016 .....</b>	<b>78</b>
<b>4. Bibliografía .....</b>	<b>83</b>

## INTRODUCCIÓN

En México ante la limitada cultura evaluativa y que muchas veces el resultado de una evaluación se “archiva” sin ser tomado en cuenta, a pesar de todos los cuidados que se hayan puesto en su preparación, una de las preocupaciones de los profesionales es ¿cómo evaluar la evaluación?, en la búsqueda de la calidad, la validez y la fiabilidad de la misma. Aparece entonces el término de **Metaevaluación** como nivel superior dentro de la teoría de la evaluación y como un nuevo concepto.

El concepto de Metaevaluación fue desarrollado por primera vez, por Michael Scriven, en 1969, quien elaboró una definición más completa, definiendo la Metaevaluación como: “La evaluación de las evaluaciones indirectamente, la evaluación de los evaluadores y representa una responsabilidad ética y científica cuando está de por medio el bienestar de otros”. Puede y debe hacerse en primer lugar por un evaluador (a) de su propio trabajo; no obstante que la credibilidad puede ser pobre, los resultados son considerados como un plus de la validación del proceso. Aun cuando los resultados de la auto-evaluación no son confiables, es deseable el uso de evaluadores independientes para realizar la Metaevaluación.

Es menester, mencionar que la Metaevaluación, del mismo modo que la evaluación, es esencialmente una actividad multidisciplinaria; por tal motivo, ninguna disciplina científica puede tener la pretensión de convertirse en la ciencia evaluadora por excelencia.

Es decir, en la práctica la Evaluación tiene dos grandes adversarios: el tiempo y la subjetividad. Evaluar es un proceso complejo, difícil de complementar en un tiempo limitado y más aún si no se tiene una cultura evaluativa. Es prácticamente imposible que todo un programa pueda ser evaluado pormenorizadamente, a la perfección y que a su vez este proceso evaluativo esté exento de insuficiencias y de subjetividad de las partes implicadas, ya sea de los evaluadores o de los propios evaluados.

Por ello, la Metaevaluación toma gran relevancia e interés en la presente tesis, ya que, desde que la metaevaluación ha ido perfilándose como un tipo de investigación de indudable interés, nos conlleva a sostener la verdadera justificación de por qué es necesaria la Metaevaluación; no sólo porque existan discrepancias con los resultados alcanzados en la evaluación o en la necesidad de identificar las posibles irregularidades que se puedan haber cometido durante el proceso.

Por tanto, la presente tesis expone la realización de la Metaevaluación a la evaluación de diseño del Programa de Inclusión Social PROSPERA 2016. Partiendo de la definición etimológica, el prefijo griego “meta” indica que el estudio se efectúa como “más allá” o “después de” la evaluación; lo que da como resultado una “evaluación de la evaluación” o la “evaluación de los evaluadores”; en esta investigación se privilegia el análisis, la reflexión y la postura crítica.

Dicho lo anterior, la tesis realizada es el resultado del análisis de la evaluación de diseño, la cual fue tomando forma en el momento de apreciar los criterios seleccionados para la realización de la misma. Con base en lo anterior se formularon los siguientes cuestionamientos: ¿Los criterios establecidos en la realización de la evaluación de diseño del Programa de Inclusión Social PROSPERA 2016, fueron los apropiados en calidad, tuvieron validez interna y externa, confiabilidad, objetividad y relevancia? Y ¿Con base en qué fueron seleccionados los criterios como focalización, objetividad, codificación, para llevar a cabo la evaluación de diseño del Programa de Inclusión Social Prospera 2016?

Ahora bien, con el fin de dar respuesta a las preguntas anteriores, el objetivo de la investigación giró en torno a la identificación de los criterios establecidos en la evaluación de diseño, así como conocer el grado de calidad y comprender los resultados o productos de la evaluación.

Con base en lo anterior, para llevar a cabo la presente tesis, partiremos de tres capítulos: en el primer capítulo se retomaron los antecedentes históricos y el análisis del tema de la Metaevaluación, con el objetivo presentar sus antecedentes,

conceptualizaciones y evolución del mismo; retomando autores como Scriven (1976); Stufflebeam y Shinkfield (1987); Schwandt y Halpern (1988); Patton (1997); Díaz y Rosales (2003); García (2003); Bustelo (2003); Cardozo (2006); por mencionar algunos.

Dado que la metaevaluación aún no cobra gran relevancia, solo en algunos países como: Reino Unido, Estados Unidos, Canadá, Francia y España. En el capítulo 2 se destaca, por identificar, describir y analizar las metaevaluaciones que se han realizado de algunos de los países anteriormente mencionados. En América Latina es casi inexistente, debido a que no se ha tenido la cultura de la evaluación, sin embargo, no haciendo menos, algunos trabajos realizados, se analizarán más adelante; con el objetivo de tener un sustento teórico-práctico, para generar un proceso de aprendizaje que, de pauta a racionalizar el proceso de investigación, principalmente los lineamientos de la evaluación (insumos, criterios, procedimientos, métodos, resultados e impactos).

Finalmente, en el capítulo 3 estriba en abordar específicamente la Metaevaluación de la evaluación de diseño del Programa de Inclusión Social PROSPERA 2016, en la que se analizan los criterios siendo este, el punto central de la investigación, ya que se realiza la en este apartado, se desglosan puntos importantes, desde la descripción de programa, hasta los lineamientos establecidos para la realización de la evaluación de diseño. Cabe mencionar, que la Metaevaluación de la evaluación de programas y políticas públicas es una realidad que, aunque tímidamente, ha ido construyendo su relevancia positivamente, ha generado tal interés hasta tal punto de reclamar la atención de profesionales y de académicos.

## **CAPÍTULO 1: CONSIDERACIONES TEÓRICAS DE LA METAEVALUACIÓN**

El presente capítulo discutirá el concepto de Metaevaluación; entendida como la evaluación de la evaluación de programas y proyectos sociales. En él, se trata de esclarecer el contenido y la definición de ese concepto.

En los subtemas que abarca el capítulo en comento, se explican las obras llamadas clásicas, así como algunos trabajos y diversas nociones propuestas por los precursores de la disciplina de la evaluación.

### **1.1. Antecedentes teóricos de la Metaevaluación**

Entre las obras llamadas clásicas, escritas por los autores considerados como principales exponentes de las teorías de la evaluación, son pocos los casos en que se incluyen capítulos o algún tratamiento orgánico del tema de la Metaevaluación. Existen claras excepciones, es el caso del autor Michael Scriven<sup>1</sup>, a quien se le atribuye la paternidad del nombre, así como de muchos otros términos del lenguaje en la disciplina.

La obra de Michael Scriven (1969) constituye un aporte trascendente. No en vano le debemos el aporte funcional del concepto de Metaevaluación; su primera formulación se ubica a finales de 1968 cuando publicó un artículo para contestar a las preguntas de cómo se evaluaban los instrumentos de evaluación. La premisa de su reflexión se resume en la siguiente idea: “puesto que la evaluación se aplica a todos los esfuerzos humanos, entonces, también la propia evaluación debe ser evaluada” (Sime, 1998: 200). Por lo tanto, los evaluadores deben buscar que sus propuestas evaluativas o evaluaciones finalizadas estén sujetas a una evaluación competente.

---

<sup>1</sup> Reineke y Welch (1986) señalan que la primera obra donde se acuñó el término fue en el artículo de 1975, difundido en Western Michigan University bajo el título “Evaluation bias and its control”. Empero, Patricia Rogers (1997), refiere otro artículo de Scriven, titulado: “An Introduction to meta-evaluation, publicado en Educational Product Report y que fue publicado en 1969, por lo que sería este el primer documento de referencia.

Es importante destacar que la visión de Scriven (1969), representa un paso muy importante en la búsqueda de legitimar la necesidad de un análisis crítico de la evaluación. En su opinión, el metaevaluador puede y debe servirse no solo de aquellos estándares que han sido específicamente formulados para la Metaevaluación, sino también de aquellos otros que han sido concebidos para la realización de evaluaciones.

De esta forma, Scriven al introducir el concepto como la evaluación de la evaluación; afirmaba que “los evaluadores tienen la obligación profesional de que las evaluaciones propuestas o finalizadas estén sujetas a una evaluación competente” (Citado en Stufflebeam y Shinkfield 1987). Es decir, su base racional es que la evaluación debe ser evaluada antes de su realización y de la difusión final del informe.

Es a partir de este punto que el término Metaevaluación, así como la actividad meta evaluadora mediante el aporte de autores como Stufflebeam (1974), House (1980), Schwandt y Halpern (1988), Scriven (1991) y Stufflebeam y Shinkfield (1993), entre otros empezaron a discutir formalmente sus procedimientos y criterios.

Daniel Stufflebeam en su libro Meta-evaluación (1974) plantea que una buena evaluación requiere que sus propios logros sean evaluados, ya que muchas cosas pueden ser y frecuentemente son hechas mal en el trabajo evaluativo. Entre los posibles problemas a verificar en la evaluación son: la parcialidad, el error técnico, dificultades administrativas o el mal uso de la misma. Tales comprobaciones son necesarias tanto para el perfeccionamiento en beneficio de las actividades evaluativas que se están llevando a cabo, como para evaluar la valía de los logros de la evaluación culminada (Dopico, s/a: 6).

En consecuencia, Stufflebeam define como objetivo de la Metaevaluación: el “asegurar la calidad de los servicios evaluativos, evitar o enfrentarse a las prácticas ilegales o a los servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la

empresa evaluativa” (1974). Además, al describir el surgimiento de la Metaevaluación Stufflebeam plantea que Leon Lessinger (1970), Malcom Provus (1973), Richard Saligman (1973) y otros han discutido el concepto bajo la etiqueta de auditoría educativa.

En relación con todo lo anterior, Stufflebeam define la Metaevaluación, más específicamente, como un proceso para describir y enjuiciar la actividad evaluativa y en contraposición al conjunto de ideas concernientes a qué constituye una buena evaluación.

Por estas razones, para Stufflebeam en el desarrollo de una metodología para la Metaevaluación es importante tener presente un conjunto de criterios apropiados. Ahora bien, los criterios para juzgar la investigación son sugeridos según Dopico (s/a), en artículos de Campbell y Stanley (1963); Gephart, Ingle (1967); y Bracht y Glass (1968). Específicamente estos autores están de acuerdo en que la investigación debe llegar a conclusiones que sean válidas interna y externamente, así como auténticas y generalizables. Aunque las evaluaciones deben también responsabilizarse por estos estándares técnicos, ellos no son suficientes para juzgar los hallazgos de la evaluación.

Otros tres estándares de suficiencia técnica, utilidad y costo/efectividad han sido explicados por Guba y Stufflebeam (1970), y por el Phi Delta Kappa Study Committee en forma de 11 criterios específicos: para medir la suficiencia técnica se plantean cuatro criterios: validación interna y externa, fiabilidad y objetividad; y para medir la utilidad se plantean siete criterios: pertinencia, trascendencia, alcance, credibilidad, oportunidad y la accesibilidad, costo/ efectividad.

Contrario a lo que dice Stufflebeam, sugiere Schwandt y Halpern (1988), que no resulta posible establecer unos estándares universalmente válidos para la realización de la Metaevaluación, sino que dichos estándares dependerán de lo que el investigador estime que es la calidad en evaluación.

Por su parte, según Dopico (2008) la American Evaluation Association (AEA) presento, en 1995, los “Principios guías para los evaluadores” con la intención de desarrollar un conjunto de principios generales y no de estándares específicos.

No obstante, Stufflebeam distingue dos roles en el metaevaluador: el formativo, que guía la evaluación, examina su proceso, asiste a los evaluadores a tomar decisiones acerca de la planificación y al desarrollo e interpretación de los resultados; el segundo rol es el sumativo, que atestigua la fortaleza o debilidad de una evaluación, teniendo así, la realización de un examen dirigido a concluir con un informe sobre su mérito (Rosales: s/a, 2).

Posteriormente, surge el libro de Ernest House, *Evaluación, Ética y Poder* (Madrid, 1994)<sup>2</sup>. House alterna el tratamiento de temas de la teoría general de la evaluación con la temática propia de la Metaevaluación. Por ejemplo, en su presentación de un conjunto de enfoques principales (los que a veces sin mayor distinción llama paradigmas, y otras, modelos básicos) pareciera moverse a lo largo de un continuo que en uno de los extremos concentra opciones “duras”, como el análisis de sistemas, el modelo de los objetivos y en de la decisión, que enfocan aspectos como la causalidad, la eficacia y la eficiencia, la programabilidad y la cuantificación de variables; mientras que en el otro extremo se ubican las opciones “suaves”, centradas en la comprensión, que presentan descripciones complejas, holísticas y que implican gran cantidad de variables interactivas”.

Respecto a House, Cardozo menciona que “las primeras opciones (duras) podemos vincularlas a definiciones de Metaevaluación de contenido práctico y operativo; mientras que las segundas (suaves) se asocian a la búsqueda más profunda de sentido y explicación de los procesos evaluativos” (2006: 133).

---

<sup>2</sup> Publicado originalmente en inglés en 1980, cuenta con una sección dedicada específicamente a la Metaevaluación e incluye apéndices en los que realiza ejercicios de aplicación.

Debe finalmente mencionarse que House identifica como determinantes de la evaluación, o lo que él llama “elementos principales para comprender los enfoques”, a la ética, la epistemología y la política (Dopico, 2008: 173).

Otra obra importante y que provee un aporte significativo en cuanto a enfoque, como en relación con el instrumental que se desarrolla, es el libro de Schwandt y Halpern (1988), *Linking Auditing and Metaevaluation*. Estos autores destacan la importancia de los criterios de rigor y la capacidad de convencimiento de las evaluaciones. Luis Fernando Díaz al respecto nos dice que para ellos “la Metaevaluación debe satisfacer esos criterios en la doble dimensión; no solo debe verificar su existencia al realizar auditorías de evaluaciones naturalísticas, sino que deben encontrarse también en la base de la calidad de la Metaevaluación” (2001: 173).

En el procedimiento que proponen Schwandt y Halpern para el auditoraje de evaluaciones naturalísticas se encuentran dos modelos paralelos:

Uno es un algoritmo, mediante el que se establece una secuencia de fases y subfases generales del proceso y la identificación correspondiente de las diversas actividades que deben desarrollar el evaluador y metaevaluador, tanto independientemente, como actuando juntos. En esta secuencia se les otorga una importancia significativa a las etapas de preparación y de planeación son centrales de determinación de la auditabilidad (evaluabilidad de la evaluación), negociación y la formalización del contrato de Metaevaluación y la planificación orientada a la obtención de verosimilitud y credibilidad. El otro modelo, que se mueve en forma paralela con el primero, es una lista de comprobación o inventario de preguntas críticas para el metaevaluador.

Es decir, el primero se trata de la revisión por el metaevaluador, solo o conjuntamente con el evaluador del programa, de las etapas desarrolladas en el proceso, entre las que resaltan las de preparación del trabajo: análisis de su factibilidad, negociación del contrato, planeación actividades, etcétera; el segundo,

se refiere al proceso mismo de investigación y consta de una lista de comprobación para garantizar las condiciones de confiabilidad y validez de los resultados (Cardozo: 2006,133).

Por otro lado, en el diseño del trabajo de Schwandt y Halpern, no parece existir la voluntad de integrar elementos epistemológicos y metodológicos originados en distintos paradigmas; es así como mucho del lenguaje utilizado, que es propio de la corriente positivista, y varias de las condiciones de confiabilidad y validez, sobre las que se establecen numerosos controles en la lista de comprobación, parecen originarse en el clima intelectual dominante en la época en que el trabajo se escribe. Lo que corresponde, además, con el propósito de utilizarlo como un “manual” para aplicarle auditorías o metaevaluaciones a las evaluaciones (Díaz: 2001, 173).

Desde el punto de vista de los autores en comento, la Metaevaluación es considerada como metáfora de la auditoría fiscal, es decir, ha de servir para mejorar la calidad de las evaluaciones; más concretamente, la Metaevaluación debe en primer término, comprobar si la metodología de la evaluación es o no la apropiada, en segundo término, si la información producida es correcta. Para ello, ha de ser relevante, fiable, suficiente, representativa y haber sido recabada a tiempo. Solo si se cumplen estos criterios, las interpretaciones, juicios y recomendaciones serán adecuados.

De entre todas estas contribuciones hemos de destacar, en primer lugar, la obra de Michael Scriven (1991). En ella discute la posibilidad y las implicaciones de realizar Metaevaluación tanto formativa como sumativa, y presenta, la idea fundamental de “la Metaevaluación como el imperativo profesional” de la disciplina, en tanto que implica el reconocimiento de que la responsabilidad ante los resultados no solo es exigible a otras actividades humanas, sino, primero a la evaluación.

No obstante, para él la verdadera Metaevaluación es una fase, usualmente terminal, del trabajo o proyecto de evaluación, que consiste en la aplicación de listas

de comprobación para construir un juicio sobre el resultado de la evaluación original. Es decir, para él, el metaevaluador puede servirse no sólo de aquellos estándares que han sido específicamente formulados para la Metaevaluación, sino también de aquellos otros que han sido concebidos para la realización de evaluaciones. De esta consideración surgieron sus dos propuestas: La Meta-Evaluation Checklist y la Key Evaluation Checklist.

La Meta-Evaluation Checklist (MEC), incluye una serie de ítems, que el autor obtiene de diversas fuentes, sobre el deber ser del estudio evaluativo. En “la MEC se considera que la evaluación ha de explicitar y justificar los estándares de valor utilizados y ser comprensiva, comprensible, creíble, ética factible, suficientemente precisa desde el punto de vista técnico, pero sin que ello haga olvidar cuestiones como la relevancia de la información proporcionada, sensible desde el punto de vista político, esto es, que atienda y responda a las necesidades de las distintas audiencias, oportuna, segura, válida y rentable en términos de coste-eficacia” (García: 2003,37).

La Key Evaluation Checklist (KEC), en principio elaborada como referencia para la realización de evaluaciones, puede también resultar de utilidad al metaevaluador. La KEC, que el autor ha revisado y actualizado recientemente (Scriven, 2001), no debe ser considerada como una lista desiderata, sino más bien como un catálogo de cuestiones que deben ser analizadas en un proceso metaevaluativo. Dichas cuestiones se refieren al diseño del estudio, el personal de la evaluación, su calendario, la identificación de las audiencias, la determinación de los costes y recursos, la valoración del proceso de investigación y de sus resultados, las alternativas existentes a la evaluación, la posibilidad de generalización de ésta y, por último, la significación global de la investigación evaluativa.

Asimismo, en la KEC se contempla la posibilidad de que, en el informe, el metaevaluador incluya recomendaciones sobre el uso o el destino de la evaluación.

Es decir, en esta lista Scriven propone que la evaluación debe ser evaluada preferiblemente antes de su realización y de la difusión final del informe. La evaluación externa es deseable, pero primero el evaluador principal debe aplicar la lista de indicadores a la propia evaluación. Los resultados de la Metaevaluación deben ser utilizados formativamente, pero también pueden ser incorporados al informe o transmitidos (resumidamente) al cliente y a otras audiencias adecuadas (Dopico: 2008, 37).

Una lista similar había sido preparada en la Western Michigan University en 1975, por Daniel Stufflebeam (Meta Evaluation). Una versión depurada de las listas de comprobación de Scriven y de Stufflebeam es presentada por Stufflebeam y Shinkfield, en su libro Evaluación Sistemática (1993). La cual “diferencia entre el cliente que encarga la evaluación y sus consumidores, y que agrega nuevos elementos como la función que cumple el evaluado, el sistema que utiliza para distribuir sus servicios, el significado o síntesis de todo lo analizado y la Metaevaluación que puede incorporarse al propio informe de la evaluación” (Cardozo, 2006: 135).

Sin embargo, otra de las referencias obligadas en el campo de la Metaevaluación es la obra de Daniel Stufflebeam (1999). La propuesta más reciente de este autor, la denominada Program Evaluations Metaevaluation Checklist, constituye una adaptación al ámbito metaevaluativo de los estándares de evaluación de programas del Joint Committee: utilidad, viabilidad, pertinencia y rigor. Su finalidad es servir de guía para verificar si un proceso evaluativo se atiene a estos estándares, que son considerados el paradigma de la calidad, el deber ser de la evaluación.

Para resumir, es importante mencionar que la Metaevaluación también ha sido reconocida por especialistas más recientes como Monnier (1995), Patton (1997), Weiss (1998), Bañón 2003), Cardozo (2006) Scriven (2007), Dopico (2008).

En definitiva, tanto las obras recientes como las consideradas clásicas siguen la idea fundamental de Scriven. Para Dopico (2008) como Cardozo (2006) coinciden en que la Metaevaluación persigue el perfeccionamiento de la calidad de la evaluación, mide la veracidad, viabilidad, objetividad y fiabilidad de la información recopilada y generada en dicho proceso, lo que implica evaluar el proceso y el sistema de evaluación en cuestión.

Finalmente, Dopico agrega una nueva utilidad práctica de la Metaevaluación: dar respuesta a las discrepancias surgidas durante la evaluación, cuando la parte evaluada considera que el proceso no fue justo y transparente y tiene quejas sobre cualquiera de sus aspectos y técnicas empleados, la ética y profesionalidad del evaluador, etcétera.

## **1.2 Concepto y objetivos de Metaevaluación**

En este subtema se trata de explicar y fundamentar, así como comparar distintos conceptos y definiciones de Metaevaluación. Esta clarificación es conveniente porque las diversas acepciones del concepto, podrían conducir a aplicaciones y a resultados divergentes.

El concepto de Metaevaluación fue desarrollado por primera vez, por Michael Scriven, en 1969, quien elaboró una definición más completa posteriormente, definiendo la Metaevaluación como: “La evaluación de las evaluaciones indirectamente, la evaluación de los evaluadores y representa una responsabilidad ética y científica cuando está de por medio el bienestar de otros. Puede y debe hacerse en primer lugar por un evaluador (a) de su propio trabajo; no obstante que la credibilidad puede ser pobre, los resultados son considerados como un plus de la validación del proceso. Aun cuando los resultados de la auto-evaluación no son confiables, es deseable el uso de evaluadores independientes para realizar la Metaevaluación”. Scriven (1991, citado en S/N, 2005).

Scriven (1976) subraya así que esta fase, usualmente terminal, del trabajo de evaluación responde a la responsabilidad exigible a todas las actividades humanas, misma que adquiere especial importancia en el caso de la evaluación, por lo que la Metaevaluación representa el imperativo profesional de la disciplina, imprescindible para controlar el sesgo del evaluador.

Daniel Stufflebeam compartiendo este punto de vista, plantea que una buena evaluación requiere que sus propios logros sean evaluados, ya que muchas cosas pueden ser y frecuentemente son hechas mal en el trabajo evaluativo. Entre los posibles problemas a verificar en la evaluación están: la parcialidad, el error técnico, dificultades administrativas o el mal uso de la misma. Tales comprobaciones son necesarias tanto para el perfeccionamiento o en beneficio de las actividades evaluativas que se están llevando a cabo, como para evaluar la valía de los logros de la evaluación culminada.

Por consiguiente, Stufflebeam define como objetivo de la Metaevaluación el “asegurar la calidad de los servicios evaluativos, evitar o enfrentarse a las prácticas ilegales o a los servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la empresa evaluativa” (1974: s/p).

Asimismo, con todo lo anterior, Stufflebeam (1974) define la Metaevaluación, más específicamente, como “un proceso para describir y enjuiciar la actividad evaluativa y en contraposición al conjunto de ideas concernientes a qué constituye una buena evaluación. Por eso plantea que al entenderse la Metaevaluación como una forma de evaluación, su conceptualización debe ser coherente con la conceptualización de la evaluación” (Dopico, 2007:38).

Años más tarde Stufflebeam (1981) define a la Metaevaluación, como: el “proceso de diseñar, obtener y realizar la información descriptiva y el enjuiciamiento acerca de la utilidad, del valor, práctico y de la adecuación técnica de una

evaluación, para guiarla y hacer público el informe de sus puntos débiles y fuertes” (Rosales, s/a: 25).

De esta manera Stufflebeam da a la Metaevaluación un carácter profesional, lo cual implica enjuiciar el valor y el mérito de lo que representa y hace la profesión del evaluador: servicios de evaluación, utilización de las evaluaciones, investigación evaluativa y desarrollo organizacional.

Por otro lado, para Cook & Gruder (1978) es la evaluación de la evaluación, y las recomendaciones de cuando hacerla varían desde hacerlas después de cada evaluación hasta hacerlas periódicamente. La Metaevaluación se puede aplicar en cualquier momento de la evaluación, desde la planificación de la misma y los métodos, hasta una evaluación completa que requiera la opinión de evaluadores externos.

Similar a Cook y Gruder, el Joint Committee on Standards for Educational Evaluation (1981) define la Metaevaluación como la “evaluación de la evaluación”, pero también habla del examen y verificación independientes de la calidad de un plan de evaluación, de la suficiencia con que fue puesto en práctica, de la precisión de los resultados y de la validez de sus conclusiones, lo que supone una descripción más detallada de las implicaciones de la evaluación.

Por otro lado, las definiciones más actuales, que sin duda alguna siguen la huella dejada por Scriven se encuentra Pacheco, Juan Francisco (2010) al hablar de Metaevaluación se refiere indistintamente a dos conceptos similares pero que difieren en su aplicación; por un lado uno más práctico aunque restringido, que la circunscribe a una etapa terminal del proceso de evaluación, como recopilación de las evaluaciones ex-ante y concomitantes realizadas a un proyecto, política o programa a fin de entregar insumos al agente crítico que toma las decisiones. Así para Sanín (1999) la Metaevaluación debe concentrarse en la reformulación de políticas y de instrumentos de desarrollo derivados de los análisis de programas, y reforzar la verificación del impacto de los programas de mayor alcance.

Asimismo, tenemos una concepción más amplia de la Metaevaluación, que tiene como fin analizar y estudiar los procesos de “evaluación”, es decir, la efectividad y la realización del proceso y cuál ha sido o será su impacto real como insumo para él o los agentes críticos que deciden el futuro de una política o de un programa. De esta manera, para Bustelo (2001) el foco de este proceso evaluativo está en los métodos empleados para evaluar y no en los resultados que ésta arroja, y para Díaz y Rosales (2003) la Metaevaluación tiene como objetivo la explicación de una evaluación como un proceso social complejo.

Sin embargo, tanto Luis Fernando y Rosales Rotsay (2003) en la realización de procesos de evaluación mencionan que se utilizan dos acepciones de Metaevaluación. Una es la actividad o conjunto de actividades que consisten en la elaboración de listas de comprobación y en la aplicación directa de estas listas a un caso específico para, con base en ellas, construir, un juicio sobre la evaluación. La otra se refiere a dar sentido y explicar el significado del conjunto de actividades humanas que denominamos evaluación, en una situación concreta. La Metaevaluación no es un proceso esencialmente técnico, sino que tiene naturaleza política y ética, es un proceso social de aprendizaje, que genera a su vez otros procesos de aprendizaje no menos importantes.

Ambas acepciones son útiles y posiblemente necesarias. Formalmente la segunda es inclusiva y contiene, como una alternativa temporal y de grado o alcance a la primera. La Metaevaluación no solo transcurre por los hechos y datos “evaluados”, sino que trasciende la evaluación.

En resumen, la Metaevaluación ha sido y es comúnmente conceptualizada como aquella investigación sistemática cuyo objetivo es emitir un juicio acerca de la calidad, la relevancia, la pertinencia o los méritos de una evaluación que incluye el análisis, la interpretación, la explicación. En este sentido, la evaluación se convierte en el objeto de estudio del metaevaluador (Stufflebeam, 1981; Scriven 1969, 1991, citados en García, 2000). Por su parte Santos (1998) plantea que la Metaevaluación

consiste en analizar el proceso de evaluación. En este sentido, Rosales (1990) y Stufflebeam y Shinkfield (1987) coinciden al plantear que es necesaria una Metaevaluación para poder atribuirle un valor a cualquier proceso de evaluación.

La importancia de la Metaevaluación radica en que a través de ella puedan salir a la luz, las posibles fallas, abusos, incongruencias, desajustes y problemas que se presentan durante la evaluación. Con esto se propicia una reflexión profunda que vaya más allá de confirmar si existe validez o no en el proceso, de la calidad de los datos o de la rigurosidad de los métodos que se han utilizado para obtener información.

De acuerdo con los planteamientos de Santos (1996) un proceso riguroso de Metaevaluación no solo permitirá valorar de manera sistemática los resultados, sino que permitirá tomar decisiones eficaces para mejorar el planteamiento, la dinámica y los modelos de evaluación.

“El proceso de evaluación en cualquiera de sus vertientes, puede tener riesgos, deficiencias y un uso inadecuado. Por tal motivo, es imprescindible establecer criterios que permitan evaluar los mecanismos de evaluación” (Santos, 1996: 13).

En este contexto, la Metaevaluación funciona como garantía de la calidad de la evaluación y puede llevarse a cabo paralelamente a la evaluación, aunque también puede realizarse una vez finalizada ésta. Según House (1981, 1993; citado en García 2000) la relevancia que ha adquirido en los últimos años la práctica evaluativa requiere, que independientemente del momento en que se lleve a cabo la Metaevaluación, los protagonistas puedan emitir sus opiniones por escrito o de forma oral, de manera que sea posible descubrir aquellos problemas, dificultades o fallas que hayan condicionado el proceso.

Finalmente, se puede afirmar que, de acuerdo a la literatura consultada, se constata que una de las preocupaciones de los investigadores en evaluación es la

necesidad de evaluar la evaluación en la búsqueda de su calidad, entendiendo esta como la demostración de su validez, fiabilidad, objetividad y su credibilidad.

### **1.3 Instrumentos de la Metaevaluación**

Para llevar a cabo una Metaevaluación es preciso tener una serie de instrumentos que sean la base teórica y práctica; que cumplan con el objetivo particular de dar pauta a facilitar el entendimiento del proyecto o programa evaluado. Por lo que retomamos a diferentes autores para darle dirección y sentido a lo que hoy llamamos instrumentos de la Metaevaluación.

Scriven (1991) retomado por Cardozo (2003) propone la aplicación de su Key Evaluation Checklist, compuesta de las siguientes 13 dimensiones o aspectos a revisar, como instrumento de evaluación que también puede ser usado en la Metaevaluación:

1. **Descripción:** diseño, equipo evaluador, calendario del estudio, fuentes de información, etcétera.
2. **Antecedentes y contexto:** responsable de la solicitud de evaluación, razones, evaluaciones previas, etcétera.
3. **Consumidor:** clientes y destinatarios efectivos y potenciales.
4. **Recursos:** insumos que pueden limitar el diseño y ejecución de la evaluación.
5. **Valores:** criterios de mérito en función de las necesidades del cliente y las normas profesionales.
6. **Proceso:** desarrollo de la evaluación y problemas encontrados.
7. **Resultados:** informes y sus efectos previsibles por el evaluador, así como impactos del proceso de evaluación.
8. **Costos:** monetarios y no monetarios, así como sus costos alternativos.
9. **Comparaciones:** utilización de juicios validados, el azar, etcétera.
10. **Posibilidades de generalizaciones:** posibles usos en diseños futuros, otros programas, otros lugares, etcétera.

11. **Relevancia:** síntesis de todo lo anterior que permite valorar el mérito de la evaluación asignada.

12. **Recomendaciones:** se pueden o no hacer en torno a la continuidad del programa y sus posibles ajustes.

### 13. Informe evaluador.

De esta forma, retomando las dimensiones anteriores por Scriven, es evidente que las considera como un instrumento base para llevar a cabo la Metaevaluación, las cuales son las primeras bases sólidas que van a retomar y mejorar más adelante. Díaz y Rosales (2003) que retoman a autores, como Stufflebeam y Shinkfield (1993) de forma que amplían la lista anterior agregando puntos importantes:

1. **Descripción:** ¿Qué hay que evaluar? El evaluando (agencia o agente ejecutivo del proyecto u objeto por evaluar), descrito tan objetivamente como sea posible. ¿Cuáles son sus componentes? ¿Cuáles son sus relaciones?
2. **El cliente:** ¿Quién ha encargado la evaluación? El cliente puede o no ser el iniciador de la investigación evaluativa. La solicitud puede originarse en el fabricante, la agencia patrocinadora, o el padrino legislativo, que puede o no ser su inventor.
3. **Antecedentes y contexto del evaluando y la evaluación:** Incluye la identificación de los mantenedores; la presunta naturaleza del evaluando; las expectativas acerca de la evaluación; el tipo deseado de evaluación (formativa, sumativa, u otro).
4. **Recursos (Sistemas de apoyo o valoración de los efectivos) disponibles para la utilización del evaluando y para la utilización de los evaluadores:** No se trata de lo que se utiliza. Incluyendo el dinero, la habilidad, la experiencia previa, la tecnología y las consideraciones sobre la flexibilidad. Esto es lo que define el nivel de viabilidad.

5. **Función:** ¿Qué hace el evaluando (*sic*)? Distinguir entre lo que se supone que hace función o papel propuesto o declarado y lo que de hecho hace; ambas cosas pueden incluirse en la descripción, pero, por lo general, es mejor tratarlas aparte. ¿Existen dimensiones, aspectos o componentes obvios de estas funciones?
6. **Sistema de distribución:** ¿Cómo llega al mercado el evaluando? ¿Cómo se mantiene (en servicio)? ¿Cómo se perfecciona? ¿Cómo son sus usuarios? ¿Cómo se lleva a cabo su realización, control y perfeccionamiento?
7. **El consumidor:** ¿Quién utiliza o recibe los resultados? Distinguir las poblaciones o consumidores escogidos, mercado deseado de las poblaciones de consumidores directamente afectadas de una manera real o potencial el verdadero mercado o consumidores, receptores, o clientes del evaluando. Sin embargo, debemos considerarlos a la hora de investigar los efectos totales y describirlos como parte de los afectados, impactados o implicados.
8. **Las necesidades y valores de los afectados y potencialmente impactados:** esto incluye tanto los deseos como las necesidades y los valores, así como las normas sobre el mérito y los ideales juzgados o supuestos; las metas definidas del programa, cuando se lleve a cabo una evaluación basada en metas, y las necesidades de los otros afectados. La importancia relativa de estas consideraciones, a menudo opuestas, depende de las consideraciones éticas y funcionales.
9. **Normas:** ¿existen normas preexistentes y objetivamente valoradas, acerca del mérito o el valor, que deban aplicarse? ¿Procede alguna de ellas del cliente y el consumidor, de la función y de las necesidades-valores?
10. **Coherencia entre necesidades, valores y normas:** determinar si las metas pueden ser valoradas como apropiadas, legales, éticas, etcétera.
11. **El proceso:** ¿Qué limitaciones, costes, beneficios se aplican a la operación normal del evaluando? Un proceso administrativo tiene limitaciones especialmente significativas en lo que se refiere al “grado de realización”, esto

es hasta qué punto la operación real cumple las estipulaciones del programa o las suposiciones del patrocinador acerca de su operación.

12. **Resultados:** ¿Qué efectos, deseados, o no, se han producido? Resulta útil una matriz de efectos: población afectada por tipo de efecto (cognitivo, afectivo, psicomotor, sanitario, social, ambiental). Esto puede cruzarse con tiempo de reacción (inmediata, al final del tratamiento, en otro momento), y con duración por cada componente o dimensión (si se quiere la evaluación analítica).
13. **Posibilidad de generalizaciones a otras personas, lugares, tiempos, versiones:** Otras personas significan tanto el personal como los receptores. Esto también puede denominarse posibilidad de distribución, posibilidad de venta, exportación o modificación.
14. **Costes:** Totales y por componentes. En dinero, psicólogos, de personal; costo inicial, incluyendo preparación, mantenimiento y perfeccionamiento; costos directos e indirectos, inmediatos y diferidos.
15. **Comparaciones con opciones alternativas, incluyendo las opciones reconocidas y las no reconocidas, las disponibles y las que se pueden crear:** Competidores críticos (en términos de costo/efectividad); normalmente, incluyen a aquellos que producen efectos similares o mejores por un precio menor, y mejores efectos por un coste adicional razonable.
16. **Significado:** es una síntesis de todo lo anterior. La validación del procedimiento es, a menudo, una de tareas evaluativas, más dificultosas. Normalmente, no puede dejarse solo al cliente que, por lo general, está mal equipado en experiencia y objetividad.
17. **Recomendaciones:** pueden ser o no requeridas, y pueden o no desprenderse de la evaluación. Aunque sean requeridas, puede que no sea posible proporcionar ninguna (los recursos disponibles para la evaluación resultan cruciales aquí).
18. **Informe:** el vocabulario, la extensión, el diseño, el medio, la organización y el personal para su presentación necesitan un cuidadoso examen. Debe

verificarse la confidencialidad (protección y privacidad, versus publicidad) de la información y el informe tanto como de los borradores y preliminares.

19. **La Metaevaluación:** la evaluación debe ser evaluada preferiblemente antes de su realización (evaluabilidad), y antes de la difusión final del informe.
20. **La evaluación externa es deseable, pero primero el evaluador principal debe aplicar la lista de control de indicadores a la propia evaluación:** los resultados de la Metaevaluación deben ser utilizados formativamente pero también pueden ser incorporados al informe o transmitidos (resumidamente) al cliente y otras audiencias adecuadas.

Por estas razones, tanto Stufflebeam y Shinkfield retoman parte de la estructura de la lista de Scriven. Cabe mencionar que dentro de la lista implementaron siete aspectos más precisos, que complementan a los diferentes instrumentos, para con ello, facilitar el estudio de la Metaevaluación.

Posteriormente, para enriquecer el número de instrumentos Pacheco (2010) destaca una serie de instrumentos orientados a la evaluación de impacto, evaluación de los beneficiarios y evaluaciones de resultado.

Con relación a la evaluación de impacto, su propósito general es determinar la efectividad de las políticas; las técnicas cuantitativas de recolección de información también son útiles para obtener participación en la aclaración de los objetivos de la evaluación. Revisar otros componentes como la eficacia en función de los costos; debido a que puede construir una preocupación muy especial para las autoridades responsables, en cuales recae las decisiones de reducir, expandir o reformar la intervención del evaluador.

Una vez, que los objetivos y datos disponibles estén claros, es posible comenzar con la fase de diseño del estudio de la evaluación de impacto. La elección de la metodología dependerá de las preguntas, secuencia cronológica, restricciones

presupuestarias y capacidad de implementación de la evaluación. Con el objetivo de conocer los pros y los contras de los distintos tipos de diseño.

La otra herramienta como instrumento es la de los beneficiarios, la cual está orientada a producir información que permita a la Metaevaluación retroalimentar la gestión, en otras palabras, ajustar las acciones y modificar eventualmente los rumbos a seguir a fin de hacer más eficiente el gasto social.

La recopilación de datos para el instrumento de beneficiarios es:

- 1. Estudios de casos:** ¿Recopila información que genere un recuento que puede ser descriptivo o explicativo, puede responder a las preguntas de cómo y por qué?
- 2. Grupos representativos:** Se enfoca en las conversaciones focalizadas con miembros de la población beneficiaria, el propósito es comparar las perspectivas de los beneficiarios con conceptos abstractos de los objetivos de la evaluación.
- 3. Entrevistas:** En estas, el entrevistador plantea preguntas a una o más personas y registra las respuestas de los declarantes.
- 4. Observación:** Es la observación y registro de una situación en un registro de una situación en un registro o diario, esto incluye quien participa. La observación puede ser directa o participativa.
- 5. Cuestionarios:** Elaboración de un conjunto de preguntas de encuesta cuyas respuestas se pueden codificar coherentemente
- 6. Análisis de documentos escritos:** Revisión de documentos como sus registros, bases de datos administrativos, materiales de capacitación y correspondencia.

La información que se obtenga de este proceso, será de carácter estratégico, puesto que va a contribuir en la elaboración de las distintas acciones y a la toma de decisiones relacionadas con el programa.

Finalmente, el último instrumento que menciona Pacheco es el de resultados, el cual tiene relación fundamental con la medición del cumplimiento de los productos de un determinado proyecto. Éste se expresa en los bienes y servicios que un determinado proyecto entrega a una población específica. Un claro ejemplo son los informes sobre la terminación de un proyecto.

Es importante destacar, que Pacheco (2010) enfoca sus instrumentos de Metaevaluación con una orientación dirigida a lo que es la recolección de información, en la que clarifica y especializa cada instrumento o evaluación (impacto, beneficiarios y resultados) para llevar a cabo de manera consiente la toma de decisiones.

Por otra parte, Cardozo (2006) propone un esquema de la revisión de conceptos e instrumentos, e identifica que:

- Las Metaevaluaciones son dirigidas exclusivamente al análisis del informe final; otras son orientadas a la revisión de los procesos técnico-metodológicos implicados en la investigación o aquellas que se enfocan a los anteriores en el marco de un contexto político y una determinada estructura y funcionamiento organizacional.
- Metaevaluaciones con enfoque racional, cuantitativo y de perspectiva positivista basado en el análisis estadístico de datos; con enfoque fenomenológico, holista, poco estructurado y cualitativo, fundamentado en la hermenéutica y el estudio de los aspectos políticos, organizativos o éticos, o con un enfoque mixto.
- Metaevaluaciones que utilizan diversas listas de indicadores o aspectos a revisar, propuestas por distintos expositores del tema.

#### **1.1.4 Premisas de la Metaevaluación**

Stufflebeam (1974) al definir a la metaevaluación, como un proceso para describir y enjuiciar la actividad evaluativa y en contraposición al conjunto de ideas concernientes a qué constituye una buena evaluación; plantea que, dado que la metaevaluación es una forma de evaluación, su conceptualización debe ser coherente con alguna conceptualización de la evaluación y expone así ocho premisas:

1. La evaluación es la valoración del mérito, de ese modo, la metaevaluación significa la valoración o estimación del mérito del empeño evaluativo.
2. La evaluación sirve a la toma de decisión y a la rendición de cuenta, por tanto, la Metaevaluación deberá proporcionar información proactiva de soporte (o respaldo) a las decisiones que se toman en la conducción del trabajo evaluativo y deberá proporcionar información retroactiva para ayudar a los evaluadores responsables de su trabajo evaluativo anterior. Es decir, la Metaevaluación debe ser formativa y sumativa.
3. Las evaluaciones deben valorar la finalidad, diseño, puesta en práctica y resultados, así mismo la Metaevaluación debe evaluar la importancia de los objetivos de la evaluación, la idoneidad de su diseño, la educación de la realización al diseño de la evaluación y la calidad e importancia de los resultados de la evaluación
4. La evaluación debe proporcionar una información descriptiva, crítica y además recomendaciones apropiadas. Igualmente, la Metaevaluación debe describir y juzgar el trabajo evaluativo, también debe recomendar cómo las evaluaciones pueden ser perfeccionadas y cómo los hallazgos pueden ser usados apropiadamente.
5. La evaluación debe servir a todas las personas, tanto aquellas que están involucradas en ella, como aquellas que han sido afectadas por el programa, consecuentemente la Metaevaluación debe servir a los evaluadores y a todas las personas que están interesadas en el trabajo de los evaluadores.

6. La evaluación debe ser conducida por personas internas (evaluación formativa) y externas al programa (evaluación masiva). Por tanto, se debe conducir una Metaevaluación formativa y obtener juicios externos del valor global de las actividades evaluativas realizadas.
7. La evaluación involucra el proceso de esbozar los asuntos a los que se deben dirigir, el proceso de obtención de la información necesaria y el uso de la información para la toma de decisión y la rendición de cuentas. Los metaevaluadores deben llevar a cabo tres pasos: determinar los asuntos específicos a que debe dirigirse la Metaevaluación; acopiar, organizar y analizar a la información necesaria y, finalmente, emplear la información obtenida en la toma de decisión apropiada y en la rendición de cuentas.
8. La evaluación deber ser técnicamente adecuada, útil y efectiva. La Metaevaluación debe satisfacer los mismos criterios.

A su vez, Stufflebeam (1974) determina propósitos que debe tener la metaevaluación:

- Servir a la toma de decisión y a la rendición de cuenta. La metaevaluación debe ser proactivamente para proporcionar a tiempo las recomendaciones relacionadas a cómo deben ser diseñados y conducidos los estudios evaluativos. La metaevaluación que sirve a la toma de decisión debe ser llamada metaevaluación formativa.
- Servir a las necesidades del evaluador de ser responsable por su trabajo. Este propósito requiere que la metaevaluación sea conducida retroactivamente para producir juicios públicos sobre el mérito del trabajo evaluativo culminado. La metaevaluación que sirve a la rendición de cuneta es sinónimo de metaevaluación sumativa. El análisis exhaustivo del plan revela que mucha de la información requerida de la metaevaluación sumativa está potencialmente disponible en la metaevaluación formativa. Por tanto, la

metaevaluación formativa potencialmente puede proveer una base de datos preliminar para la metaevaluación sumativa.

Al respecto, Dopico (2007) retoma la opinión de Jiménez (s/f), quien considera que la Metaevaluación tiene que ver con varias cuestiones teleológicas: el *rigor* de la evaluación en todo el proceso como su finalidad; la *mejora*, al preguntarse por el resultado de la evaluación, por la capacidad de transformación de la practica evaluada y por enlace entre la evaluación y la mejora al final y durante el desarrollo del proceso; con la *transferibilidad*, cuando la Metaevaluación se pregunta por las garantías de aplicación de las conclusiones a otras experiencias y contextos; la *ética* en que se desarrolla la evaluación, la finalidad que se le atribuye; y con el *aprendizaje*, permitir aprender de los errores y de los aciertos.

Asimismo, Dopico resalta la importancia de la transferibilidad, el cual se plantea en que cada contexto tiene su carácter concreto e irrepetible; es un proceso que permite aplicar hallazgos conseguidos a otra situación diferente o similar. Cabe señalar que las aportaciones señaladas son con el objetivo de reflexionar sobre la evaluación debido a que ayuda a comprender el fenómeno y permite elaborar pautas de intervención en situaciones nuevas.

Finalmente, en el estudio de las premisas de la Metaevaluación Cardozo (2003) retoma a Dopico (s/f) quien ofrece al igual un amplio conjunto de aspectos, a través de los cuales se puede identificar características del programa o proyecto:

- La definición de los objetivos que se persigue con la evaluación.
- La elección del momento en que realiza la evaluación.
- La duración o premura con que se realice la evaluación.
- La predeterminación de los parámetros a evaluar.
- El sistema de puntaje que se utilice.
- La pertinencia del sistema de evaluación diseñado para que sea aplicado a todos los programas.

- La identificación del sistema de evaluación diseñado para que sea aplicado a todos los programas.
- La identificación del sistema de evaluación de diseño con lo que se entiende como patrón de calidad de los programas.
- El cumplimiento por los evaluadores seleccionados con aquellas aptitudes y actitudes establecidas para ejercer como evaluadores: justeza, objetividad, honestidad, capacidad, etcétera.
- La organización del trabajo de los evaluadores.
- El nivel de participación de los interesados en el proceso de evaluación.
- El clima preponderante durante la evaluación.
- El apoyo institucional para que el proceso de evaluación se lleve a cabo en las condiciones requeridas.
- Las técnicas empleadas por los evaluadores para la recolección de la información.
- La calidad de los documentos del programa presentados por los evaluadores.
- La suficiencia de medios y recursos con que se cuenta para la realización de la evaluación.
- La utilización de dichos recursos.
- La riqueza de la información aportada por los encuestados y entrevistados.
- La calidad del informe final presentado por los evaluadores.

Retomando lo anterior, podemos decir que la lista enunciada por Dopico, son una serie de premisas con las cuales se lleva a cabo el proceso de Metaevaluación, con el objetivo de identificar por medio de su aplicación en la investigación, las debilidades o las carencias del trabajo de evaluación.

## 1.5 Tipos de Metaevaluación

Para identificar y distinguir los diferentes tipos de Metaevaluación, retomaremos la base teórica que expone María Bustelo (2003) en el cual plantea cuatro tipos básicos de Metaevaluación:

- Considera el contenido o la fase del programa evaluado: ya sea de diseño, proceso o resultados. En primer lugar, la evaluación es formativa o sumativa. Sumativa cuando la función es recapitular el proceso completo de la evaluación. La Metaevaluación es formativa cuando es utilizada como instrumento de cambio durante el proceso de la evaluación, su propósito es el de mejorar el diseño de la evaluación en proceso.
- En consideración con el contenido o la fase del programa evaluado: diseño, proceso o resultado. La Metaevaluación del diseño implica el análisis y valoración del plan y el diseño de la evaluación, enjuicia sus propósitos, objetivos, coherencia, interna, contextualización, estrategias de implementación, entre otros. La Metaevaluación de proceso analiza la implementación de la evaluación. Incluye el análisis de las dificultades en la implementación del proceso, ajustes al calendario de ejecución, adecuación a las necesidades y efectividad de los equipos de trabajos. La Metaevaluación de resultados analiza tanto la información obtenida como los resultados y efectos; ésta tipo de Metaevaluación debe de otorgar como resultado un mejor entendimiento del proceso, un mejor aprendizaje por parte de los equipos, en suma, directamente relacionado con la utilidad de los procesos.
- La evaluación ex ante (antes de la implementación del proceso) o ex postfacto (durante el proceso o inmediatamente finalizado).
- Considerando el agente que la realiza: interno o externo. La metaevaluación interna es desarrollada por los evaluadores mismo,

es decir las mismas personas que estuvieron a cargo de la evaluación, desarrollan la metaevaluación, por lo cual, son responsables del control de la calidad de los procesos. La metaevaluación externa es desarrollada como su nombre lo indica por un agente externo que no estuvo involucrado en el proceso de evaluación, lo cual como efecto le da más credibilidad al proceso y un reconocimiento externo y consolidación del mismo.

Con base en lo anterior, Bustelo hace una estructuración de los diferentes tipos de metaevaluación, teniendo pertinencia en el proceso para llevar a cabo la clasificación de las evaluaciones, debido a que agrupa criterios generales y específicos (según sea el tipo de evaluación) para darle sustento teórico, ayudar a mejorar el entendimiento del proceso, y proporcionar resultados con mayor alcance.

Bajo la misma lógica, la metaevaluación de diseño y en concordancia con Bustelo, que se enfoca en el análisis y la estructura que se tiene planeada, en secuencia menciona la metaevaluación de proceso, la cual se especializa en identificar las dificultades de la implementación, y por último la metaevaluación de resultados, que es un complemento de las anteriores, debido a que su razón de ser es ayudar a mejorar el entendimiento del proceso.

De igual forma, Cardozo (2006) define a la Metaevaluación de diseño en el sentido de que se debe analizar el contexto evaluativo, la claridad de sus objetivos, la inclusión de las variables relevantes, la justificación del esfuerzo, su costo, el tiempo disponible, su contribución potencial a la mejoría del programa, la participación de los actores involucrados, la disponibilidad de la información necesaria, la previsión de técnicas adecuadas de recolección y análisis de datos, de presentación y difusión de informes, entre otros. Asimismo, define a la Metaevaluación de proceso; analizan la puesta en práctica de la evaluación antes diseñada: ajuste al calendario, disponibilidad oportuna y uso eficiente de los

recursos, calidad por parte del equipo evaluador, la aplicación de los métodos y técnicas previstas, problemas identificados, entre otros.

Ambas concepciones de las Metaevaluaciones de diseño son similares y tienen la misma finalidad, la cual gira entorno a dar un mejor resultado, por lo que a través de su análisis se interpreta el funcionamiento de la evaluación de diseño como la forma táctica de llevar a cabo un análisis del contexto, identificando puntos claves, como lo son: el identificar la relevancia de la necesidad, al igual que la problemática social, la determinación de la población objetivo, la consistencia entre los objetivos y estrategias planteados.

Es importante destacar, que en la Metaevaluación de procesos se identifican ciertas diferencias entre ambas autoras; como también Bustelo agrega una Metaevaluación de resultados, que es herramienta clave para llevar a cabo una evaluación con más sustento teórico.

Por último, habría que decir que a las metaevaluaciones se les pueden aplicar todas las clasificaciones de evaluaciones disponibles, sin embargo, las que giran particularmente en torno al proceso, se clasifican en:

Metaevaluaciones de diseño, que deben analizar el contexto evaluativo, la claridad de sus objetivos, la inclusión de las variables relevantes, la justificación del esfuerzo, su costo, el tiempo disponible, su contribución potencial a la mejoría del programa, la participación de los actores involucrados, la disponibilidad de la información necesaria. La previsión de técnicas adecuadas de recolección y análisis de datos, de presentación y difusión de informes, etc.

Metaevaluación de proceso, que analizan la puesta en práctica de la evaluación antes diseñada: ajustes al calendario, disponibilidad oportuna y uso eficiente de recursos, calidad del equipo evaluador, aplicación de los métodos y técnicas previstas, problemas encontrados, etc.

A continuación, se muestran las propuestas y Checklist para la metaevaluación, por los diferentes autores mencionados.

**Cuadro 1.**  
**Propuestas y “Checklist” para la metaevaluación**

**Cuadro-resumen**

	<b>Dimensiones, ítems, estándares.</b>
<b>Meta-Evaluation Checklist (Scriven, 1991)</b>	<p>La evaluación debe:</p> <ul style="list-style-type: none"> <li>- -explicitar y justificar los estándares de valor realizados</li> <li>- -ser comprensible</li> <li>- -ser creíble</li> <li>- -ser ética</li> <li>- -ser factible</li> <li>- -ser precisa desde el punto de vista técnico, pero sin que se olviden cuestiones como, la relevancia de la información proporcionada.</li> <li>- ser “sensible” desde el punto de vista político, esto es, que atienda y responda a las necesidades de las distintas audiencias.</li> <li>- -ser oportuna</li> <li>- -ser segura</li> <li>- -ser válida</li> <li>- -ser rentable en términos de coste-eficacia.</li> </ul>
<b>Key Evaluation Checklist (Scriven, 1991; 2007)</b>	<ul style="list-style-type: none"> <li>- La metaevaluación debe analizar:</li> <li>- -Los antecedentes y el contexto de la evaluación</li> <li>- La definición de los términos empleados</li> <li>- Las descripciones de la evaluación que hagan sus responsables</li> <li>- Los criterios de valor</li> <li>- La identificación de las audiencias y de los stakeholders.</li> <li>- La determinación de los costes y recursos.</li> <li>- La valoración del proceso de investigación y de sus resultados.</li> <li>- Las alternativas existentes a la evaluación</li> <li>- La posibilidad de generalización de la evaluación</li> <li>- La significación global de la investigación evaluativa</li> <li>- La existencia de recomendaciones sobre el uso o el destino de la evaluación.</li> </ul>

<b>(Schwandt y Halpern, 1988)</b>	<p>La metaevaluación debe comprobar:</p> <ul style="list-style-type: none"> <li>- Si la metodología empleada en la metaevaluación es la adecuada.</li> <li>- Si las evidencias encontradas son correctas (y, por tanto relevantes, fiables, suficientes, representativas y han sido recabadas a tiempo)</li> </ul>
<b>(Program Evaluations Metaevaluations Checklist (Stufflebeam, 1999))</b>	<p>Las evaluaciones deben cumplir con los siguientes requisitos:</p> <ul style="list-style-type: none"> <li>- Utilidad</li> <li>- Viabilidad</li> <li>- Pertinencia (adecuación a las normas)</li> <li>- Rigor (precisión)</li> </ul>
<b>Guiding Principles Checklist for Evaluating Evaluations (Stufflebeam, Goodyear, Marquert Y Johnson, 2005)</b>	<p>La metaevaluación debe analizar</p> <ul style="list-style-type: none"> <li>- Si la evaluación es una investigación sistemática.</li> <li>- El grado de competencia de los evaluadores</li> <li>- La integridad y honestidad del proceso.</li> <li>- Si ha existido respeto por las personas</li> <li>- Si los evaluadores han asumido responsabilidad por el bienestar de la sociedad.</li> </ul>
<b>Evaluación democrática deliberativa (House y House,2000)</b>	<p>La metaevaluación debe comprobar si la evaluación se ha atendido a los principios de:</p> <ul style="list-style-type: none"> <li>- Inclusión</li> <li>- Deliberación</li> <li>- Diálogo</li> </ul>

Elaboración propia con base en *Metaevaluación de políticas públicas: una visión desde la ciencia política*, Ester García Sánchez. (2009)

A todo esto, García Sánchez (2012) nos menciona que, a la vista de todas estas consideraciones, parece evidente que la metaevaluación no puede quedar reducida al mero examen de la calidad técnica de una evaluación. Ciertamente es que las evaluaciones son, en lo esencial, procesos de investigación, pero también lo es que son empresas eminentemente políticas que se desarrollan en un contexto también político. De hecho, como recuerda Patton (1997), la metaevaluación es tanto más

necesaria cuanto más politizado esté en el contexto y cuanta más visibilidad e importancia hayan alcanzado las evaluaciones en él. Si bien la actividad metaevaluativa se convierte no sólo en un imperativo profesional para los evaluadores sino en una empresa de innegable interés para los académicos, los decisores y para la sociedad en su conjunto (Stufflebeam, 2001)

Si bien, pensar que es la Metaevaluación conduce, inevitablemente, a considerar los instrumentos y/o dimensiones que conviene tener en cuenta a la hora de enjuiciar la calidad de un proceso evaluativo como lo hemos visto en el presente capítulo. Sin embargo, la determinación de lo que es “calidad” en evaluación está lejos de ser un tema resuelto.

Cierto es que los últimos años se ha incrementado considerablemente el número de trabajos que se han ocupado de desentrañar aquel concepto o, en otros términos, de seleccionar los criterios conforme a los cuales se ha de juzgar una evaluación. Como es lógico, el desarrollo de la Metaevaluación es mucho más notorio en aquellos países que cuentan con una dilatada trayectoria evaluativa, como veremos a continuación. No obstante, y afortunadamente, el avance de la evaluación en otros lugares como América o España ha propiciado, en los últimos años, el interés por el tema, lo que se ha traducido tanto en la elaboración de propuestas teóricas como en la realización de estudios empíricos.

Con el fin de establecer las bases con las que se realizara la Metaevaluación, haremos a continuación un breve recorrido por las Metaevaluaciones realizadas por el mundo.

## **CAPÍTULO 2. META-EVALUACIÓN EN EL MUNDO<sup>3</sup>**

Existe muy poca investigación que sistematice y analice las experiencias de metaevaluación. En los últimos años, en países en que la cultura y la práctica de la evaluación están más desarrolladas, se ha insistido en la necesidad de realizar investigación que aproveche las evaluaciones ya realizadas. Algunos antecedentes pueden encontrarse en Reino Unido, Canadá, Bélgica, Francia y España. Casi todos refieren sus técnicas de recolección de la información y sus resultados, pero ninguno explicita el tratamiento posterior que dio a la misma para arribar a sus conclusiones.

En trabajo realizado en Canadá por Lacasse (1993) constituye el primero al que se ha tenido acceso. Se basó en la revisión de 37 informes de evaluación del período 1988-1990, complementando con entrevistas confidenciales aplicadas a funcionarios, consultores y clientes de las evaluaciones, con base en una muestra intencional en que selecciono sólo los trabajos de mayor calidad y dirigidos por personas de reconocida competencia.

Lascoumes (1998) relata una metaevaluación en Francia que dio origen a la realización del coloquio “Evaluation pluralista des politiques publiques; enjeux, pratiques et produits”. Su objetivo era enfocarse al conjunto de trabajos colectivos que permitían la materialización de la evaluación: diversidad de prácticas, herramientas, modelos teóricos y metodologías.

A pesar del avance alcanzado por países desarrollados, existe muy poca investigación que sistematice y analice las experiencias en el tema. Por ello, es necesario recabar y analizar aquellos países que se caracterizan por su cultura y práctica metaevaluativa; así como la identificación de sus técnicas y de instrumentos de las experiencias que se comentaran. Pues bien, en este capítulo se comentan algunos trabajos, así como diversas nociones propuestas en la disciplina de la Metaevaluación.

<sup>3</sup> Resumen de las experiencias reportada en Cardozo (2006 y 2009)

Ahora bien, la práctica de la Metaevaluación, se ha estado expandiendo aceleradamente en los últimos años como consecuencia de la preocupación por la eficiencia en la aplicación de los recursos que introdujeron las políticas neoliberales. Sin embargo, no es muy generalizada, y ha sido interpretada indistintamente, en dependencia de criterios, contextos y sistemas evaluativos.

Al respecto, se constataron las diferentes direcciones en las que se introduce, según algunas experiencias, la práctica metaevaluativa en el mundo, entendida como: la evaluación de la evaluación, tributando a un estudio de la forma de hacer en varios procesos de acreditación es decir un meta-análisis, experiencia de Taiwán, en 1997.

Otra experiencia más es la evaluación simultánea a un proceso de evaluación, es decir durante el proceso, es realizada en la Universidad de Veiga en Brasil. Por otro lado, el estudio de modelos evaluativos institucionales, en Cataluña, España y en EE.UU.; las “auditorias” con un informe público para determinar los procedimientos evaluativos son adecuados, en Gran Bretaña; el control de la existencia de mecanismos de aseguramiento de la calidad de las instituciones o del cuidado y atención con que se implementan los mismos, así como la validación de sus procesos y resultados.

Al respecto Cardozo (2010) hace mención de algunos antecedentes que pueden encontrarse en Reino Unido, España y parte de América Latina. Dichos países en comento serán parte fundamental de lo que se analiza en este capítulo.

Es importante destacar que la Metaevaluación se encuentra aún en proceso de desarrollo: pero cuenta con mayores avances en los países donde la evaluación se haya más arraigada y en las cuales casi todos, refieren sus técnicas de recolección de la información y sus resultados, pero ninguno explicita el tratamiento posterior que dio a la misma para arribar a sus conclusiones.

## **2.1 Experiencias de Metaevaluación en Gran Bretaña**

Cardozo (2006) hace una revisión especialmente de las experiencias desarrolladas en Estados Unidos, Canadá, Reino Unido, Francia y América Latina. En el cual hace una actualización de algunos resultados del seguimiento del tema en Europa basándose en los resultados presentados por el Barómetro 2008 de la evaluación de las políticas públicas en Europa (Ernst y Young, 2008); de lo cual resalta un punto importante:

- Las prácticas de la Comisión Europea constituyen un referente respetado por los países, y los programas europeos son los más evaluados.

Bustelo (2001) plantea que, a pesar del desarrollo alcanzado por la evaluación en Europa, aún hay poca investigación que sistematice y analice estas experiencias. En los países en que la cultura y la práctica de la evaluación están más desarrolladas se ha insistido, en los últimos años, en la necesidad de realizar una investigación que aproveche los estudios ya realizados. Algunos antecedentes podemos encontrarlos en Reino Unido.

A continuación, se presentan dos casos en los que se ha utilizado la Metaevaluación como técnica de constatación de la metodología, impacto y resultados de otras evaluaciones realizadas.

### **2.1.1 La Metaevaluación del Sistema de Evaluaciones de EuropeAid**

Una de las experiencias en Europa importante de mencionar, es la realizada por el Centro de Estudios de Cooperación al Desarrollo (CECOD) titulado “Análisis de los Resultados y Metaevaluación del Sistema Evaluaciones de EuropeAid”. El trabajo tuvo como objetivo realizar una revisión de los informes de evaluación externa encargados por la Unidad de Evaluaciones de EuropeAid, oficina ejecutora de las acciones programadas desde la Dirección General de Desarrollo de la Comisión Europea.

El fin que pretendió CECOD es obtener información sobre las buenas y malas prácticas detectadas en las misiones evaluadoras; el trabajo en comento constó de tres secciones para poder afrontar el propósito de la Metaevaluación. La primera sección cuenta con una descripción del sistema de evaluaciones dentro de la comisión.

En la segunda sección se realiza la síntesis de los resultados ofrecidos por los informes de evaluación, es decir, una *evaluation synthesis* más que una Metaevaluación (que se ofrece después) como bien distingue Bustelo (2002). Mientras que la Metaevaluación es una “evaluación de las evaluaciones realizadas” juzgando su calidad, la síntesis evaluativa trata de detectar el común denominador de los informes elaborados.

Para el autor Pacheco el trabajo realizado por CECOD, es una “síntesis de los resultados emanados de los informes de la evaluación rescatando los puntos débiles y fuertes de la ayuda comunitaria europea basándose en los propios informes de evaluación, agregando mayores informes de retroalimentación al ciclo evaluativo” (2010, 69).

Respecto a esa línea argumentativa Bustelo (2001) nos dice que la síntesis de los resultados sirve para aunar, resumir o incluso potenciar los resultados de varios estudios de evaluación de programas o políticas similares. Pero en realidad es un procedimiento para revisar estudios similares anteriores. Es decir, el énfasis que se presenta es hacer comparables los resultados de varios estudios, si bien la función última no será juzgar los procesos de evaluación, sino tener una base empírica para juzgar los resultados generales de los programas o políticas evaluadas.

Es importante destacar que la Metaevaluación que se realiza, cumple con una doble funcionalidad: por un lado, valora la calidad y pertinencia de los informes de evaluación con respecto al proceso metodológico en general, configurando las próximas líneas de investigación que serán ligadas con las decisiones futuras. Por

el otro lado se plantean dilemas y problemáticas propias de una dimensión más macro y general sobre la política de desarrollo europea, entendible a otras agencias de cooperación (CECOD: s/a, 3).

La Metaevaluación realizada derivó en el siguiente diagnóstico valorativo:

- La existencia de diversos tipos de evaluaciones recogidas en la mayoría de los casos: temáticas, sectoriales, de país, regionales, de instrumentos financieros o de líneas presupuestarias, formativas, sumativas, evaluaciones ex post e intermedias; dejan en evidencia el hecho de que la Unidad de Evaluaciones ha sabido utilizar los diversos tipos de evaluaciones según el caso, la pertinencia y utilidad para cada uno de ellos.
- Las recomendaciones realizadas son pertinentes en términos amplios, pero en algunos casos caen en una excesiva ambigüedad y generalización que se ve amortizada por el uso de la Ficha contradictoria, instrumento que permite contrastar los comentarios y recomendaciones emanadas del informe con los “alegatos” de la Unidad Evaluadora.
- De esta manera pese a que la mayoría de los informes poseen una extensión que promedia las 70-100 páginas, algunos presentan un tamaño excesivamente extenso, haciéndose ineficiente e ineficaz la lectura y comprensión de estos a la luz de la carga de trabajo y la dotación de expertos en cada Unidad.
- Si bien la falta de datos concretos hace imposible asignar un costo cuantitativo para comparar la eficiencia de cada misión evaluadora, es razonable suponer que EuropeAid opte por equipos de primer nivel, lo que supone un precio elevado por un producto de alta calidad.
- La falta de equipos evaluadores liderados o compuestos mayoritariamente por miembros de los mismos países receptores se constituye como un aspecto negativo, limitando la representación de los beneficiarios; se debe examinar la influencia real de las recomendaciones en las acciones de planificación.

- Otro punto de interés, pero que desborda el objetivo del estudio, es el uso que hace el Parlamento de los informes de Evaluación.
- Es necesario profundizar en la relación entre evaluación y asignación de recursos, priorizando por las acciones mejor gestionadas y que hayan logrado un mayor impacto. Es suma, poner en práctica las consecuencias de orientar la cooperación al desarrollo por resultados (Pacheco: 2010, 70).

Con la revisión de la experiencia llevada a cabo en Gran Bretaña, se puede afirmar que la Metaevaluación realizada por el CECOD sigue trabajos importantes; es el caso de la clasificación realizada por la Evaluation Research Society en donde la Metaevaluación puede incluir la síntesis de evaluaciones, e incluso ser la acepción principal, (Evaluation Research Society Standards Committee, 1982), y en la que uno de los 6 tipos de evaluación descritos es la Metaevaluación. En esta clasificación a la Metaevaluación se le otorga el significado de síntesis de evaluaciones, es decir, una forma de evaluación que reanaliza los resultados de una o varias evaluaciones para determinar qué es lo que se ha aprendido sobre una política o programa (Chelimsky, 1989).

La Metaevaluación en un sentido más crítico, puede afirmarse que cumplió con lo que debe tener la realización de la misma, según el autor Ballart (1992) quien nos dice que “La Metaevaluación consiste en el análisis de resultados de varias evaluaciones”, aunque también habla de la revisión de tipo más amplio de la investigación evaluativa desarrollada en un determinado periodo.

Es decir, al realizarla hace una revisión de los informes, en donde se utilizó la síntesis de los resultados para tener una base empírica para juzgar los resultados generales de los programas o políticas evaluadas. Es por ello que para el autor Pacheco; la Metaevaluación realizada por CECOD es la que mejor representa lo que debiese ser una Metaevaluación puesto que su objeto de estudio son precisamente las evaluaciones (CEPAL: 2010).

Sin embargo, no se tiene una metodología clara, o en su defecto bien estructurada, debido a que no plantean criterios para la investigación, la cual gira en torno a la Metaevaluación, cabe resaltar que las técnicas de recolección de información debe ser la base primordial para abordar el tema y realizar la correcta evaluación.

### **2.1.2 Metaevaluación de la Agenda de Modernización de los Gobiernos Locales del Reino (Local Government Modernization Agenda, LGMA)**

En el Reino Unido entre 1998 y 2001 el gobierno local impulsó más de 20 iniciativas de política tendientes a modernizar la agenda de dichos gobiernos, estas iniciativas se denominaron colectivamente Agenda de Modernización de los Gobiernos Locales (LGMA). Si bien muchas de estas iniciativas de política habían sido evaluadas de forma individual, teniendo como fin la verificación del estado de avance en su implementación particular, de esta manera proyectos como acuerdos de servicios Públicos Locales (Local Public Service Agreements), fueron analizados, pero no ofrecían una visión conjunta del progreso de la agenda, por tanto, se desconocía su real efecto en la modernización de los gobiernos locales.

Para explorar el impacto combinado de las distintas iniciativas en la LGMA, la oficina del Vice-primer ministro encargó una Metaevaluación del proceso conjunto. Como resultado se obtuvo: la ***Meta-Evaluation of the Local Government Modernization Agenda: Progress Report on Service Improvement in Local Government***

La Metaevaluación fue llevada a cabo por un equipo dirigido por el Centro de investigación del Gobierno Local y Regional en la Universidad de Cardiff. El objetivo de la Metaevaluación fue reunir pruebas de cada una de las evaluaciones para identificar las iniciativas políticas que han sido claves para el desarrollo de la modernización del gobierno local y destacar las sinergias y los conflictos entre las distintas políticas (Leach, Lowndes, Cowell & Downe, 2005).

El trabajo realizado consistió en la recolección de datos de los informes de evaluación (publicados e inéditos) individuales de políticas de la LGMA junto con una encuesta nacional de las autoridades locales y oficiales, así como también estudios de caso de seis autoridades locales, realizados específicamente para la Metaevaluación (CEPAL: 2010, 71).

Es importante mencionar que la Metaevaluación se basó en la combinación de datos secundarios de otros estudios, medidas de rendimiento y datos primarios recopilados por el equipo de Metaevaluación. Las fuentes de datos clave para el informe fueron:

- Documentos gubernamentales relacionados con LGMA
- Informes producidos por equipos de investigación que evaluaban las políticas de LGMA
- Un análisis inicial de la satisfacción del público
- Resultados iniciales de la encuesta de los funcionarios de las autoridades locales, realizadas por el equipo de Metaevaluación

El trabajo realizado desarrolló una “teoría del cambio”, es decir, un modelo simplificado de cómo se puede esperar que el LGMA tenga un impacto en la mejora del servicio, dicho modelo respalda la Metaevaluación llevada a cabo. El cual proporcionó un marco para identificar problemas clave y posibles áreas de cambio en las que la Metaevaluación debió enfocarse e identificar los datos que se necesitaron para recolectar y analizar.

Además, Pacheco (2010) en dicha Metaevaluación, ordenó las conclusiones extraídas de los informes de evaluación de acuerdo a los siguientes puntos:

- Mejora de servicios;
- Rendición de cuentas;
- Liderazgo de la comunidad;
- Participación de las partes interesadas;

- La confianza del público.

Pacheco (2010) al respecto, nos dice que el trabajo realizado por el Centro de Investigación local y Regional; emplea una metodología participativa en la realización de la Metaevaluación. En ella se analizan los estados de avance de los proyectos implementados y los grados de cumplimiento de las metas propuestas por esta agenda, con lo que se puede afirmar el rumbo a seguir y las decisiones que son susceptibles de tomar. No obstante, lo más importante es que se puede establecer la coherencia entre las iniciativas de política presentadas y los proyectos y acciones ejecutadas (2010, 65).

La Metaevaluación presentó un informe final presentado al Vice-primer Ministro, el cual tuvo como objetivo tomar una visión más amplia y proporcionar una evaluación general del estado actual del conocimiento sobre el impacto combinado de esas políticas en la mejora del servicio, así como el evalúa el impacto del progreso acumulativo de las políticas dentro de la LGMA sobre la cantidad y calidad de la autoridad local y su compromiso con los diferentes interesados directos (stakeholders engagement).

La Metaevaluación antes mencionada, se puede identificar que los evaluadores partieron de la recolección de informes para la identificación de las deficiencias en las evaluaciones, cabe resaltar que los evaluadores tuvieron como base, indicadores para tener un control sobre la estructura apegado a una metodología para su irrealización, debido a que aún no se plantean un método o criterios más enfocados al tema.

## **2.2 Experiencias de Metaevaluación en España**

En lo que respecta al caso español, su cultura de evaluación ha estado poco desarrollada y se ha iniciado a raíz de su incorporación a la Unión Europea. El proceso de institucionalización de la misma, comienza con el Real Decreto Número

951 de 2005 sobre Mejora de Calidad en la Administración del Estado y con los trabajos sobre el análisis del impacto normativo encomendados por Acuerdo del Consejo de Ministros (Cardozo: 47).

### **2.2.1 Metaevaluación de evaluaciones de políticas públicas de igualdad de género en España**

El siguiente caso es el de la autora María Bustelo; quien para optar el grado de Doctor realiza un trabajo titulado: “La evaluación de las políticas públicas de igualdad de Género de los Gobiernos Central y Autonómicos en España: 1995-1999.

El objeto de estudio de la tesis doctoral que presentó, es la evaluación de políticas públicas aplicado al caso concreto de las políticas de igualdad de género. Para ello, se dio a la tarea de realizar una Metaevaluación de las evaluaciones de los planes de igualdad de oportunidades entre hombres y mujeres (planes de igualdad de género), realizadas en España en los niveles de gobierno central y autonómico, en la segunda mitad de la década de los años noventa.

Es importante destacar que Bustelo (2001); le dio prioridad a la Metaevaluación como procedimiento para describir, analizar y valorar los procesos de evaluación que se han producido en relación a las políticas de igualdad de género en España. Es decir, no se plantea la Metaevaluación como una forma de controlar la calidad de las evaluaciones realizadas, sino como una forma de estudio de las políticas públicas, en su fase concreta de evaluación.

Ahora bien, más que la valoración de la calidad de las evaluaciones estudiadas, se realizó una valoración de su adecuación y oportunidad al contexto evaluativo y de la política, y de su contribución al cumplimiento de las funciones de la evaluación.

Lo que pretendió Bustelo (2001), además de lo ya mencionado, es dar una visión global sobre cómo se desarrolla la función de evaluación en el sector de políticas de igualdad de género en España.

Por esas razones, la autora optó por tener por objeto de la Metaevaluación un conjunto de varias evaluaciones en lugar de centrarse en uno o dos procesos de evaluación. Se favoreció la descripción general comparativa entre varias evaluaciones por encima de la descripción pormenorizada y en profundidad de cada proceso de evaluación.

Respecto al tipo de Metaevaluación realizada Bustelo (2001), señala lo siguiente:

- Es una **Metaevaluación sumativa**, ya que hace recapitulación de los procesos de evaluación habidos y completados en un tiempo determinado. La intención es realizar un análisis global y sumativo de dichos procesos.
- Es una **Metaevaluación post-facto** ya que los procesos de evaluación que se meta evalúan ya han ocurrido y se describen y valoran a posteriori.
- Es una **Metaevaluación externa** ya que se realiza por alguien ajeno tanto a los equipos de evaluación que realizan las diversas evaluaciones como a los organismos de igualdad que las encargan.
- Finalmente, pretende ser una **Metaevaluación de diseño**, de proceso y de resultados.

El propósito de la Metaevaluación es contribuir al estudio de la evaluación de las políticas públicas de igualdad de género y los objetivos de la misma son, por un lado, analizar la función de la evaluación en torno a dichas políticas y, por otro, proponer alternativas para la mejora de dicha función. El objeto de la Metaevaluación son los procesos de la evaluación en sí mismos, no los resultados que arrojan. Por ello Bustelo (2001) plantea tres grandes criterios de Metaevaluación:

1. ¿Cómo se han planteado las evaluaciones y cuales han sido las estrategias evaluativas planteadas?
2. ¿Cuál ha sido el papel de algunos elementos clave en la evaluación?
3. ¿Cuál ha sido la utilización de las evaluaciones?

Al respecto Bustelo (2001) para llevar a cabo el primer criterio, analizó el contexto evaluativo, la claridad o ambigüedad en los propósitos y las funciones de las evaluaciones, las estructuras institucionales de la evaluación. En el segundo se centró en el estudio de los agentes y los participantes, los momentos y los tiempos, los criterios e indicadores, así como los procedimientos e instrumentos de recopilación y análisis de la información utilizados. Por último, el tercer criterio, se analizó la pertinencia y utilidad de la información producida, la comunicación de los resultados y el impacto de los resultados de la evaluación en la política.

Bustelo (2001), en su trabajo aplicado a la Metaevaluación; analiza la evaluación de once planes y programas en el tema. En el cual utiliza técnicas de recopilación y análisis documental; entrevistas abiertas semiestructuradas aplicadas a los responsables de los organismos involucrados; y entrevistas estructuradas, personales y telefónicas, realizadas a los responsables técnicos de la evaluación en los mismos organismos.

En la Metaevaluación de las políticas de igualdad de género de los gobiernos central y autonómico de España realizada por Bustelo (2001), no se encuentran criterios que aludan explícitamente a la política o a la ética de las evaluaciones. La autora agrupa los criterios de Metaevaluación en tres grandes categorías:

1. Como se plantean las evaluaciones y cuáles son las estrategias evaluativas empleadas.
2. El papel de algunos elementos clave en la evaluación
3. La utilización de las evaluaciones.

En la primera se encuadran criterios, tales como: el contexto evaluativo, la claridad o ambigüedad en los propósitos y las funciones de las evaluaciones, las estructuras institucionales de la evaluación, los tipos de evaluación por los que se opta y los recursos empleados en los procesos de evaluación. En la segunda se incluyen el estudio de agentes y participantes, los momentos y los tiempos, los criterios e indicadores, así como los procedimientos e instrumentos de recopilación y análisis de datos utilizados.

La tercera y última engloba el examen de la pertinencia y utilidad de la información producida, la comunicación de los resultados y el impacto de los resultados de la evaluación en la política pública (García: 2011, 9).

Bustelo (2004) nos aclara que “la Metaevaluación no se enfoca a los resultados de las evaluaciones sino a los procesos implicados en la realización de ellas, revisa, entre otros, si cuentan con evidencia suficiente, si la metodología ha sido adecuada, si las interpretaciones, juicio y recomendaciones fueron lógicamente diseñados (Cardozo/ Evalúa 6).

Es decir, no se plantea la Metaevaluación como una forma de controlar la calidad de las evaluaciones realizadas, sino como una forma de estudio de las políticas públicas, en su fase concreta de evaluación. Si bien, más que realizar una valoración de calidad de las evaluaciones estudiadas, se realizó una valoración de su adecuación y oportunidad al contexto evaluativo y de la política, y de su contribución al cumplimiento de las funciones de la evaluación.

El propósito del estudio metaevaluativo que se presentó fue contribuir al estudio de la evaluación de las políticas públicas de igualdad de género y los objetivos de la misma fueron:

1. Analizar la función de evaluación en torno a dichas políticas
2. Proponer alternativas para la mejora de dicha función

El objeto de la Metaevaluación son los procesos de evaluación en sí mismos, no los resultados que arrojan. El objeto de análisis de y valoración del trabajo se centró en la evaluación de los planes, como máxima y principal expresión de las políticas de igualdad de género

La función que ejerce la Metaevaluación, no es juzgar los procesos concretos sino la función de la evaluación en las políticas públicas de igualdad de género, así como contribuir al estudio de la evaluación.

Las estrategias planteadas en el análisis del planteamiento de una evaluación de la Metaevaluación como metodología son:

- Analizar el contexto evaluativo
- Analizar los propósitos de cada uno de los procesos de evaluación, su claridad y coherencia
- Estudiar el tipo de estructura a las que se asigna la función de evaluación
- Se señalan los tipos de evaluación realizados y los recursos empleados en las mismas.

Con el análisis de los dos casos que realizó Bustelo demuestra que el utilizar técnicas de recopilación y análisis documental; así como entrevistas abiertas semiestructuradas aplicadas a los responsables de los organismos involucrados y entrevistas estructuradas personales y telefónicas, realizadas a los responsables técnicos de la evaluación en los mismos organismos resalta la fiabilidad así como coherencia que se tuvo en la evaluación; teniendo así una Metaevaluación más completa en el aspecto de la aplicación de técnicas de investigación, así como su metodología, ya que privilegia al método comparativo.

### **2.2.2 La evaluación de Programas de Reforma Educativa en España: Tres estudios de caso desde un enfoque de Metaevaluación.**

Su interés se centró en el análisis del binomio actores decisiones en dichos procesos de evaluación.

Existe otro trabajo de Metaevaluación, de enfoque cualitativo, efectuado por Ester García (2003 y 2004). La investigación llevada a cabo, se circunscribió en el estudio de tres de las evaluaciones de programas o proyectos educativos de reforma más significativos en España en el periodo 1982- 1992.

La selección de los casos se realizó a partir de tres criterios convergentes:

- Que los programas educativos que fueron objeto de la evaluación perteneciesen a niveles de enseñanza no universitaria;
- Que el diseño de dichos programas hubiese correspondido en exclusiva a la administración educativa central;
- Que las evaluaciones de los programas hubiesen sido financiadas por la administración educativa central.

Así, la selección quedó finalmente reducida en los siguientes tres casos:

1. Evaluación externa de la reforma experimental de las Enseñanzas Medias (1984 –1992)
2. Evaluación del programa de integración escolar de alumnos con necesidades educativas especiales (1985-1989)
3. Evaluación del proyecto Atenea para la introducción de la informática en la escuela (1985-1990)

Los cuatro objetivos en los que se basó la investigación fueron:

1. Reconstruir y analizar, es decir, tomaron como referente el proceso de institucionalización de la evaluación.

2. Identificar los criterios relevantes para la clasificación de los actores de las evaluaciones, así como los factores que condicionan las decisiones en las que éstos participan, todo ello a efectos de elaborar un esquema de análisis aplicable a todos y cada uno de los casos seleccionados
3. Determinar, a partir de las clasificaciones propuestas, el papel desempeñado por los distintos actores implicados en cada uno de dichos procesos de evaluación, así como el peso específico de cada uno de ellos; en definitiva, analizar la distribución del poder en el seno de las evaluaciones.
4. Analizar las actitudes y el comportamiento de los diferentes actores en cada una de las fases del proceso de evaluación y especialmente, sus percepciones sobre el impacto de las evaluaciones.

Desde la perspectiva de García Ester, desde que Michael Scriven acuñara el término a finales de los años sesenta, la Metaevaluación ha ido perfilándose debido sobre todo a las contribuciones pluralistas como un tipo de investigación de indudable interés para determinar la calidad de las evaluaciones. Calidad que, como apuntan autores como House (1980, 1993), Scriven (1969, 1991), Patton (1997, 2000) o, incluso, el propio Joint Committee on Standards for Educational Evaluation (1994) no puede medirse únicamente en función del rigor técnico de la evaluación sino también a partir de otros criterios como el grado de pluralismo del proceso evaluativo, la utilidad de los resultados, la actitud de los poderes públicos con respecto a la propiedad de los estudios o la medida en que estos pueden contribuir a la profundización de la democracia.

La investigación realizada por García, descansa en el primer criterio de la argumentación, es decir, tomó como punto de partida la tesis pluralista que subraya el carácter político de las evaluaciones.

Pues bien, el enfoque naturalista pluralista, sostiene a través de House, Stake y Scriven, la importancia del componente político en una evaluación, destacando los aspectos éticos en el proceso de investigación evaluativa. Dichos aspectos

pasaron a convertirse en el estándar con el cual se debe medir el valor de la evaluación, puesto que esta ha de realizarse conforme a requerimientos de carácter ético como el de respeto a la pluralidad de opiniones existentes, a la confidencialidad de los informes, entre otros. Es por ello que Ester García en la Metaevaluación realizada se centra en valorar el grado en que la investigación evaluativa cumple con esos requisitos, por tanto, el significado de utilidad que se emplea se modificara conforme varíe el enfoque.

Al llevar a cabo dicha investigación constato en primer lugar, que la Metaevaluación es la culminación analítica de los procesos de evaluación, así como confirmo la relevancia de la Metaevaluación y que constituye una gran herramienta para conocer los estilos de evaluar de los poderes públicos.

### **2.3 Experiencias de Metaevaluación en América Latina**

En América Latina la Metaevaluación es un reto a enfrentar, pero sobretodo es una oportunidad para inmiscuirse dentro de los procesos para enjuiciar las evaluaciones para buscar y encontrar esa fórmula que toda política pública busca para llegar a la perfección, al igual que su retroalimentación.

Sin embargo, su aplicación práctica ha sido escasa, especialmente en los países sudamericanos, donde no existe una cultura de proyectos consolidada y, por ende, una cultura de evaluación.

#### **2.3.1 Metaevaluación en Costa Rica**

El trabajo realizado en Costa Rica (2002), específicamente para la Universidad de Costa Rica (Facultad de Ciencias Económicas “Escuela de Estadística, maestría profesional en población y salud” por Helen Larissa Centeno Monge. Trata de realizar una Metaevaluación del programa antidrogas y Salud Sexual y Reproductiva para Adolescentes, que se implementó en escuelas públicas de la ciudad de Kingston, Jamaica, durante los años de 1992 y 2001 (Centeno: 2002, 3).

El nombre original del programa es “Addiction Alert Adolescent Drug Prevention Program” el cual tiene como propósitos primordiales el de promover un estilo de vida saludable, libre de drogas y prevenir conductas autodestructivas y/o antisociales incluyendo el abuso en el uso de drogas y aumentar el conocimiento sobre temas de Salud Sexual y Reproductiva entre la Juventud de alto riesgo que habita en el área metropolitana de Kingston Jamaica.

El objetivo general fue realizar una Metaevaluación del programa Antidrogas y Salud Sexual y Reproductiva que se implementó en los adolescentes de las escuelas Públicas de Kingston, Jamaica en los años de 1999-2001.

En el trabajo realizado por Centeno, la Metaevaluación persigue evaluar la planificación e implementación del programa, así como el de evaluar las estrategias de análisis estadístico con las que se trabajaron los datos y las interpretaciones que se les dieron a los resultados obtenidos durante el proceso de Evaluación de Impacto que se realizó en la Universidad de Tulane, New Orleans.

La contextualización de dicha Evaluación que Centeno retomo para su evaluación es que dicha investigación se realizó como respuesta a la inquietud de la Organización patrocinadora, de saber el impacto logrado con su programa y así promover su continuidad.

Dicho proyecto pretendió llegar a una contribución metodológica basada en la Teoría Psicométrica, que fortaleciera los conocimientos de futuros evaluadores y los incentive a darle importancia tanto a los análisis cuantitativos como cualitativos dentro de una Evaluación de impacto.

La Metaevaluación realizada en el proyecto contemplo cuatro etapas:

1. Diseño del programa
2. Instrumento
3. Análisis de datos llevada a cabo en la Universidad de Tulane
4. Valoración de la interpretación

A partir de la Metaevaluación realizada por Centeno, podemos resaltar la utilización de su metodología y la aplicación de criterios coherentes a la evaluación realizada, también partimos de la idea clara y precisa sobre dos planteamientos básicos el ¿Qué se iba a evaluar? Y el ¿Cómo se iba a evaluar?; para clarificar su presente Metaevaluación, Centeno presentó un pequeño listado para identificar las etapas por las cuales iba a realizar el proceso del mismo; iniciando por el diseño del programa, instrumento, análisis de datos y la valoración de la interpretación, cada etapa lleva un orden y un proceso diferente.

Autores como **Rossi** concuerdan con la idea de Centeno sobre la importancia de la “valoración de necesidades”, que va orientado al diseño del programa; las siguientes etapas marcadas, giran en torno a criterios como pertinencia, utilidad, costo/efectividad, por mencionar algunos de los que plantean autores como Stufflebeam, y que con base en ellos es posible llevar un análisis más especializado. Si bien, esta Metaevaluación es una de las más completas en el sentido de la utilización con ya lo mencionábamos en su metodología, instrumentos de investigación, criterios y ese ciclo evaluativo para generar un análisis bien sustentado, que en efecto tiene como resultado un enriquecimiento y aprendizaje sustancial en materia de Metaevaluación.

## **2.4 Experiencias de Metaevaluación en México**

En el contexto mexicano, la normatividad ni siquiera la menciona. A nivel federal se ha planteado en diversas oportunidades el interés de realizar metaevaluaciones con asesoría de algún organismo internacional que analicen, valoren y contribuyan a mejorar la calidad de las evaluaciones realizadas. El CONEVAL no ha realizado ninguna; en cambio existen varias efectuadas en forma independiente a las evaluaciones federales. Es importante señalar que, en 2013 la SEDESOL publicó trabajos que denominó como tales para cubrir las evaluaciones realizadas a 20 programas en 2007-2012.

### Metaevaluaciones 2007-2012 coordinadas por el SEDESOL

Programa	Institución Evaluadora
Desarrollo Humano oportunidades	Universidad A. De Chiapas
Abasto Social de Leche	Universidad A. De Chiapas
Abasto Rural	Universidad A. De Chiapas
Atención a Jornaleros Agrícolas	Universidad A. De Chiapas
70 y más	Universidad A. De Chiapas
Coinversión Social	Universidad A. De Chiapas
Apoyo Alimentario	Universidad A. De Chiapas
Apoyo a las Instancias de Mujeres en las Entidades Federativas	Universidad A. De Chiapas
Apoyo a avenidas en Pobreza matrimonial para Regularizar Asentamientos Humanos Irregulares	El Colegio de la Frontera Norte
Desarrollo de Zonas Prioritarias	El Colegio de la Frontera Norte
Ahorro y Subsidio para la Vivienda Tu Casa	El Colegio de la Frontera Norte
Vivienda Rural	El Colegio de la Frontera Norte
Hábitat	El Colegio Mexiquense
3 x 1 para Migrantes	El Colegio Mexiquense
Rescate de Espacios Públicos	El Colegio Mexiquense
Prevención de Riesgos en Asentamientos Humanos	El Colegio Mexiquense
Opciones Productivas	CIDE
Empleo Temporal	CIDE
Estancias Infantiles para Apoyar a madres Trabajadoras	CIDE

Fomento Nacional para el Fomento de las Artesanías.	CIDE
---	------

Fuente: Elaboración propia con base en SEDESOL (2018)

En este apartado se busca recopilar y analizar las Metaevaluaciones realizadas en México de los programas de desarrollo social a cargo de la Secretaría de Desarrollo Social (SEDESOL). Por lo que Cardozo (2006) hace un diagnóstico sobre el estado de la evaluación de las políticas y programas en función, que cubran los objetivos revisados (describir, analizar y valorar los procesos de evaluación y la calidad de sus resultados) y desarrolle los determinantes de carácter organizativo. Para ello analizó los informes finales de evaluación, con un enfoque multidisciplinario, tanto cuantitativo como cualitativo.

#### **2.4.1 Metaevaluación del Fondo Nacional para el Fomento de las Artesanías 2012**

La Metaevaluación 2007-2012 del Fondo Nacional para el Fomento de las Artesanías (FONART) realizada por el Centro de Investigación y Docencia Económicas (CIDE). Tuvo por objeto analizar los avances y logros del programa en términos de evaluación y monitoreo, durante el periodo 2007-2012. Asimismo, en la meta evaluación se buscó identificar las áreas de oportunidad del programa para mejorar su desempeño en lo referente al monitoreo y evaluación, en el periodo 2013-2018” (Cejudo, 2012: 2).

La Metaevaluación en comento propuso la implementación de una evaluación de indicadores con el fin de verificar que los indicadores de la MIR del FONART en efecto estén capturando la información para la cual fueron diseñados. La evaluación complementaria propuesta al igual busca generar evidencia empírica para conocer los atributos de diseño con los que debe contar el programa para ser más eficaz (Cejudo, 2012).

Además, a partir de la implementación de la evaluación del proceso, se busca conocer los aspectos que deben ser mejorados en la operación del programa. En la

agenda de evaluación se incorporó también una evaluación de impacto con miras a conocer los efectos del programa en el ingreso de los beneficiarios, en el desarrollo de sus condiciones productivas y en la preservación de la actividad artesanal.

En materia de monitoreo, se realizó un análisis histórico (2007-2012) de los cambios que ha sufrido la Matriz de Indicadores, identificó que la MIR redujo el número de indicadores a lo largo del periodo de análisis.

La evaluación analizó los siguientes criterios:

1. Pertinencia
2. Cobertura
3. Utilidad
4. Credibilidad

Con base en la Metaevaluación del Fondo Nacional para el Fomento de las Artesanías del 2012, se tuvo por objetivo analizar los avances y logros en materia de evaluación y monitoreo durante el año antes mencionado, así como identificar áreas de oportunidad en dichos temas, que permitan a los programas obtener mejores resultados. Si bien la Metaevaluación incluye también un análisis de viabilidad y pertinencia de la evaluación de impacto para el FONART.

Si nos basamos en lo postulado por los diversos autores mencionados, en el anterior capítulo, la Metaevaluación logra identificar una gran variedad de particularidades, en cuanto al diseño de la evaluación realizada, que pueden ser útiles para mejorar el diseño de la evaluación.

En general la Metaevaluación encontró principales fortalezas, así como una serie de recomendaciones a programa evaluado a fin de mejorar la operación y planeación. Es decir, al basarse su evaluación en un trabajo de gabinete en el que se revisaron los resultados finales de la MIR, el análisis de la cobertura del FONART, así como de su presupuesto y evaluaciones previas al programa, cumple con las

premisas que Stufflebeam establece para la adecuada realización de una Metaevaluación. Sin embargo, se considera deseable revalorar su método y hacer explícito el uso de los instrumentos establecidos. Es importante hacer hincapié en este último punto, pues al no hacer uso de los mismos que son la base teórica y práctica para cumplir con el objetivo particular de dar pauta a facilitar el entendimiento del proyecto o programa evaluado, es imposible determinar si estos tuvieron un buen desempeño, calidad y confiabilidad en la realización de la misma.

#### **2.4.2 Metaevaluación del Programa de Abasto Rural (PAR) a cargo de DICONSA 2007-2012**

La Universidad Autónoma de Chiapas realizó la meta- evaluación del Programa de Abasto Rural, quien se planteó como objetivo el analizar los avances y logros en materia de evaluación y monitoreo durante el periodo 2007- 2012, así como identificar áreas de oportunidad que permitan al programa obtener mejores resultados en próxima administración federal.

En síntesis, se enfoca en la metodología y los principales resultados; identifica las fortalezas, oportunidades, debilidades y amenazas más relevantes, y enuncia las recomendaciones que los evaluadores derivaron en su momento de las evaluaciones.

La evaluación de consistencia y resultados del Programa de Abasto Rural a cargo de DICONSA, tuvo como fin retroalimentar el diseño y la gestión del programa a partir del examen de la consistencia del programa en cuanto a:

- Diseño
- Planeación estratégica
- Cobertura, focalización
- Operación
- Percepción de la población objetivo y
- Resultados.

## **Análisis**

Es importante destacar que existen algunas debilidades adicionales en estos trabajos organizados por la SEDESOL: *a)* puesto que un metaevaluador analiza críticamente evaluaciones previas nunca puede coincidir con la persona responsable de las últimas; pero en este caso eso sucede; *b)* todas las metaevaluaciones fueron concentradas en cuatro equipos de trabajo, sin mediar convocatoria abierta o por invitación (se trata de adjudicaciones directas); *c)* una Metaevaluación suele requerir información adicional a la documental como la realización de entrevistas, especialmente a funcionarios y evaluadores, de cuya realización no hay indicios; y *d)* no hay ninguna justificación evidente para la diversidad de montos pagados por Metaevaluación a cada equipo.

Ahora bien, las metaevaluaciones de FONAR Y PAR como ambas instituciones las denominan contaron con una valoración técnica de las evaluaciones realizadas de 2007 a 2012. Empero, solo sirvió como una herramienta para generar una agenda de evaluación pertinente para los próximos ejercicios fiscales. Las cuales incluyen un análisis de viabilidad y pertinencia de la evaluación de impacto para el Programa de Abasto Rural. El objetivo de ambas “metaevaluaciones” fue analizar los avances y logros en materia de evaluación y monitoreo, así como identificar áreas de oportunidad para los programas.

“Más que una Metaevaluación se trata de una síntesis analítica del contenido de evaluaciones previas que incluyen: un análisis FODA, hallazgos relevantes, propuesta de una agenda de evaluaciones futuras, mejoras a las MIR, conclusiones y recomendaciones y darles seguimiento. Pero ninguna analiza la calidad de los trabajos de evaluación realizados: su suficiencia, rigurosidad, claridad, viabilidad de sus recomendaciones” (Cardozo, 2015: 198).

Es decir, no se trata de una Metaevaluación en el sentido amplio de la palabra ya que la misma tiene por objeto de análisis a las evaluaciones (no a los programas evaluados) en sus aspectos metodológicos, organizativos y éticos, la primera parte del objetivo es válida como guía para su realización, pero no la segunda.

En efecto, existen algunos ejemplos de “metaevaluaciones” realizadas de manera independiente a las evaluaciones de los programas federales de SEDESOL. Sin embargo, estos trabajos no necesariamente constituyen una verdadera Metaevaluación más bien constituyen análisis que intentan valorar el impacto de los programas.

#### **2.4.4 Evalúa DF 2008**

En lo que va del nuevo siglo, la institucionalización formal de la evaluación ha avanzado en México a nivel federal, especialmente con la creación del Consejo Nacional de Evaluación de la Política Social (Coneval) en 2006, sin que se haya podido generar una verdadera cultura que la institucionalice a nivel sustantivo. Por su parte el gobierno del Distrito Federal (GDF) careció de un sistema de evaluación que sustentara sus decisiones hasta que en 2007 se creó Evalúa DF, como un organismo público descentralizado, encargado de manera exclusiva e independiente de la evaluación externa de la política y los programas sociales, por cuenta propia o a través de terceros, con la particularidad de que sus recomendaciones tienen carácter vinculatorio, a diferencia de Coneval. (Cardozo, 2010: 1)

Evalúa DF inicio actividades en junio de 2008 con la participación de seis consejeros académicos. En 2008-2009 llevó a cabo 13 evaluaciones de diseño y operación de programas sociales, cuya Metaevaluación 2009 vamos analizar en este subtema del capítulo 2.

En el año 2007 se decidió la creación de Evalúa DF, encargado de la evaluación de políticas y programas sociales a nivel de la capital (no incluye otros sectores). En el caso de Evalúa DF, éste llevó a cabo distintos tipos de metaevaluaciones considerando la importancia de reflexionar sobre la calidad de lo realizado: a) evaluación de cada informe por uno o dos consejeros, discusión y decisión

colectiva sobre su aprobación, solicitud de correcciones o rechazo, así como similar proceso posterior a la aprobación para diseñar las recomendaciones que serían vinculatorias; b) evaluación en conjunto de evaluaciones correspondientes a las 13 evaluaciones de programas realizadas en 2009.

En 2009, el Consejo de Evaluación del Desarrollo Social del Distrito Federal, Evalúa DF, llevo a cabo un ejercicio de Metaevaluación para analizar y valorar críticamente el trabajo realizado por los evaluadores contratados durante el año anterior (2008). Decidió someter a examen su propio trabajo en materia de evaluación de programas de desarrollo social durante (2008).

En este aspecto, la política social del Distrito Federal buscó promover: la generación de actividades económicas para la creación de empleos dignos; materializar y ampliar los derechos económicos, sociales, culturales y ambientales de las y los ciudadanos; incrementar la cobertura de los programas; así como realizar acciones que permitan garantizar el acceso equitativo y la calidad de los servicios, bienes e infraestructura públicos a todos sus habitantes.

Para llevar a cabo el proceso de Metaevaluación se retomaron elementos valorados, con base en los criterios generales de investigación:

1. Objetividad
2. Validez
3. Confiabilidad
4. Representatividad
5. Consistencia, etc.

Y algunos criterios específicos de la Metaevaluación como:

1. Pertinencia
2. Autonomía
3. Credibilidad
4. Utilidad

5. Participación

6. Calidad.

Cabe mencionar que, en el proceso de evaluaciones externas, el consejo tomó la decisión de hacer 13 evaluaciones de diseño y operación a programas sociales, ocho de los cuales son responsabilidad del gobierno central del DF y otros cinco más de las delegaciones políticas. Con base en lo anterior se establecieron al igual criterios, de los cuales destacan:

1. Presupuesto asignado al programa
2. Cobertura del programa.
3. Importancia del programa para solucionar los principales problemas sociales de la población del DF.
4. Carácter emblemático del programa para la población del DF y su gobierno.
5. Grado de maduración del programa.

Aunado a esto, los criterios generales para dictaminar la calidad de los informes intermedios fueron:

1. El grado de avance del informe final en comparación con el cronograma propuesto por el equipo evaluador.
2. La coherencia lógica de las aseveraciones mencionadas en los documentos.
3. La suficiencia de información utilizada para fundamentar dichas aseveraciones.
4. La claridad, la importancia y la relación de los problemas encontrados en el análisis.
5. La pertinencia y actualización de la bibliografía utilizada para el análisis teórico de los problemas que los programas intentan resolver.
6. El apego general a los términos de referencia.

El consejo también emitió comentarios y sugerencias para ser incorporados en el informe final, pero respetando plenamente las valoraciones hechas por el

equipo evaluador. Para aprobar el informe final se utilizaron además otros cuatro criterios:

1. El apego a la propuesta técnica presentada por los evaluadores.
2. La fundamentación del análisis hecho.
3. La valoración de la gestión analizada.
4. La propuesta de recomendaciones.

Para la realización de los informes se retomaron fuentes documentales o secundarias pertinentes y actuales (leyes, reglamentos, decretos, planes, programas, estadísticas, presupuestos, libros, artículos, informes de investigación, ponencias, etc.) y entrevistas como fuente primaria. Las últimas utilizaron guiones semiestructurados y estructurados. Sólo uno menciona explícitamente la observación como método de recolección de información.

En relación con los recursos empleados, cinco evaluaciones presentan una cobertura completa y clara del punto, basada en un análisis de la evolución de los mismo, especialmente los financieros, en periodos variables entre dos y seis años, su relación con las metas, nivel de suficiencia y necesidades inmediatas, porcentaje del gasto social en el presupuesto delegacional; complementado con la disponibilidad de recursos humanos y materiales. Información que uno de los quipos no habría obtenido.

Respecto a la valoración de diseño, la mitad de los informes recuperan los hallazgos que han ido identificando a lo largo del análisis y realizan una valoración adecuada del conjunto.

En relación con la valoración de la operación, los informes recuperan adecuadamente los principales hallazgos realizados y obtienen conclusiones claras y bien fundamentadas, si bien algunos tienen deficiencia en la valoración.

En síntesis, el Consejo de Evaluación del Desarrollo Social del Distrito Federal consideró importante llevar a cabo este tipo de Metaevaluación, para

analizar y valorar críticamente el trabajo realizado por los evaluadores contratados. Cabe resaltar que este trabajo constituye la primera Metaevaluación y autoanálisis de la también primera generación de evaluaciones del Desarrollo Social en el Gobierno del Distrito Federal, al igual pretende institucionalizarse de manera periódica.

Se puede afirmar, que cuenta con una congruencia y estructura clara, dado que los informes de EVALUA DF se caracterizaron por un alto grado de homogeneidad en la estructura de su contenido y por la sistematización del trabajo realizado. Esto se debe a gran medida que los términos de referencia se apegaron a una serie de lineamientos disponibles y al libro de orientación con el que cuenta el Consejo. Es por ello que Evalúa DF ha hecho hincapié en que sus evaluaciones se lleven a cabo de manera sistemática, haciendo uso de los métodos de investigación social, emitiendo juicios de valor sobre el desempeño analizado, recomendaciones precisas y bien argumentadas.

Finalmente se pudo concluir que, en el caso de Evalúa DF, la Metaevaluación de 2009, logró, en primer lugar, mejorar su organización y funcionamiento en relación con la precisión de TR, tiempos de trabajo, selección de evaluadores, métodos y técnicas de evaluación, entre otras para contar con mejores evaluaciones y recomendaciones. En segundo lugar, impulsó la posterior realización de al menos una evaluación de impacto y dio mayor impulso al desarrollo de un sistema de seguimiento de las recomendaciones vinculatorias generadas. Adicionalmente al mostrar su contribución al logro de una mejor calidad, reforzó la intención de generar una nueva Metaevaluación externa en 2010.

Más allá del impacto directo provocado por la Metaevaluación en 2009 en la organización y funcionamiento de Evalúa DF, estos trabajos han inducido un lento interés de los investigadores y un entusiasmo considerablemente mayor entre los estudiantes universitarios, lo que en términos de marco teórico se resume en aprendizaje traducido en modificación de las evaluaciones y avances conceptuales

y metodológicos de quienes se encuentran en procesos de formación. Sin embargo, su institucionalización se vislumbra aún lejana.

Por ello, como universitarios y dando pauta a que la Metaevaluación tome un rumbo diferente. Damos pie al análisis de la Evaluación propuesta a evaluar, contemplando los postulados de los autores vistos en los capítulos anteriores, para que las únicas experiencias de Metaevaluación que existen, pasen de ser solo aplicaciones de modelos teóricos.

## **CAPÍTULO 3. ANÁLISIS DE LA EVALUACIÓN DE DISEÑO DEL PROGRAMA DE INCLUSIÓN SOCIAL “PROSPERA 2016”**

A la vista de todas las consideraciones hechas en capítulos anteriores, parece evidente que la Metaevaluación no puede quedar reducida al mero examen de la calidad técnica de una evaluación.

De este modo, nuestra propuesta de Metaevaluación ha pretendido aproximarse más a la visión de Scriven como a la de Stufflebeam, es decir a un enfoque metaevaluativo cásico y convencional que gira en torno a estándares previamente establecidos. Por tanto, hemos tomado como referente a la Metaevaluación formativa y sumativa, partiendo de la Key Evaluation Checklist (Scriven: 2001) y las 8 premisas que plantea Stufflebeam (1974).

### **3.1 Programa de Inclusión Social PROSPERA**

Prospera Programa de Inclusión Social es un programa conceptualizado e instrumentado por el Gobierno Federal en 1997, como su principal estrategia para incidir en el fenómeno de la pobreza, la articulación del programa para alcanzar el desarrollo social integral de la población es una decisión del Estado mexicano para atender dicha situación. El cual a través de dos esquemas: con corresponsabilidad y sin corresponsabilidad, otorga recursos a casi 7 millones de familias mexicanas para fortalecer su alimentación, salud, educación; asimismo vincula a las personas beneficiarias con proyectos productivos, opciones laborales y servicios financieros (SEDESOL: 2016).

#### **3.1.1 Descripción General del Programa**

El programa en comento inició el 8 de agosto de 1997 con una cobertura de 300 mil familias en el ámbito rural. Posteriormente a principios de 2002 PROGRESA<sup>3</sup>

---

<sup>3</sup> Se usa el nombre PROGRESA “Programa de Educación, Salud y Alimentación” como antecedente del programa PROSPERA, vigente de 1997 a 2002.

atendía a 2.4 millones de hogares, de los cuales, dos terceras partes eran comunidades indígenas.

A mediados del 2002, el Programa se transforma a Oportunidades y aumenta su cobertura a los 32 estados del país, alcanzando así los 4.2 millones de hogares atendidos. Los beneficios para las familias aumentaron, ya que las becas educativas, que al principio únicamente para educación básica, se ampliaron a educación media superior.

El 5 de Septiembre de 2014, a través de un Decreto Presidencial, Oportunidades se fortalece y se transforma en PROSPERA Programa de Inclusión Social, cuyo objetivo es articular y coordinar la oferta institucional de programas y acciones de política social, cuyo objetivo es articular y coordinar la oferta institucional de programas y acciones de política social, incluyendo aquellas relacionadas con el fomento productivo, la generación de ingresos, el bienestar económico, la inclusión financiera y laboral, educación, alimentación y salud, dirigidas a la población en situación de pobreza, bajo esquemas de corresponsabilidad que les permitan a las familias mejorar sus condiciones de vida y asegurar el disfrute de sus derechos sociales y el acceso al desarrollo social con igualdad de oportunidades (GOB: 2018).

Su objeto establecido en el decreto que lo creó, es el de articular y coordinar la oferta institucional de programas y acciones de política social, incluyendo aquellas relacionadas con el fomento productivo, generación de ingresos, bienestar económico, inclusión financiera y laboral, educación, alimentación y salud, dirigida a la población que se encuentre en situación de pobreza extrema, bajo esquemas de corresponsabilidad.

En 2015, en las Reglas de Operación de PROSPERA<sup>4</sup>, primeras que se publicaron después de su creación, el Programa amplía su objetivo general respecto

---

<sup>4</sup> Para 2016, el artículo cuarto transitorio del ACUERDO por el que se emiten las Reglas de Operación de PROSPERA Programa de Inclusión Social, establece que: Derivado del presupuesto de egresos de la

al de su antecesor, Oportunidades para incluir el acceso a otras dimensiones de bienestar que, en términos de su decreto de creación, se refieren al fomento productivo, generación de ingresos, bienestar económico, inclusión financiera laboral, sin embargo no refleja esto en sus conceptos de apoyo, pues permanecen los que establecían en 2014: Seguridad alimentaria, acceso a la educación y cobertura de salud.

### **3.2 Evaluación de Diseño**

Para proceder a la Metaevaluación de la evaluación<sup>5</sup> de diseño del Programa de Inclusión Social Prospera 2016 es preciso; conocer, explicar y fundamentar en que consiste la evaluación de diseño, así como la conceptualización que el Consejo Nacional de Evaluación (CONEVAL) de la Política de Desarrollo Social fundamenta ya que al ser la institución responsable de emitir los lineamientos de Evaluación y al realizar la evaluación en comento es fundamental presentarlo.

Para Osuna (2000) la evaluación de diseño se enfoca en la planificación de la política pública, al igual que evidencia la necesidad del cual es relevante la presencia del programa, así como la racionalidad en la cual se desarrolla; también se debe de priorizar la importancia sobre la tesitura en la que el diseño de una política pública debe de ir de la mano de la evaluación, para la evolución constante.

Es decir, la evaluación de diseño es una herramienta de planificación, en el que su diseño es el objeto de análisis y se debe de considerar como una de las evaluaciones más importantes debido a que permite tomar decisiones para mejorar

---

Federación para el Ejercicio Fiscal 2016, considera la fusión del programa presupuestario S118 PROGRAMA DE Apoyo Alimentario con el S072 PROSPERA Programa de Inclusión Social, las familias beneficiarias que hasta el Ejercicio Fiscal 2015 habían sido atendidas por el Programa de Apoyo Alimentario, se integrarán en el Esquema de Apoyos sin Corresponsabilidad y las que habían sido atendidas por el PROSPERA Programa de Inclusión Social, se integrarán en el Esquema de Apoyos con corresponsabilidad (RO: 2016).

<sup>5</sup> “La evaluación permite conocer el nivel de desempeño de políticas y programas (resultados e impactos), explicarlo tomando en consideración las características de su diseño e instrumentación, pero también el contexto político, social y económico concreto en que se inserta, y valorarlo a la luz de un conjunto de criterios como la pertinencia, la congruencia, la eficiencia, la calidad, la equidad entre otros” (Cardozo, 2015: 13).

la lógica interna, lo que da pauta a contribuir a identificar las fallas y reajustar el programa por lo cual fue creado.

Otro autor que hace referencia a este tipo de evaluación es Bustelo (2001) quien reconoce que este tipo de evaluaciones es el que menos se practica. Probablemente la extendida asunción de que la evaluación debía, básicamente, centrarse en valorar el alcance de los objetivos incidió en una menor producción y difusión teórica en materia de evaluación conceptual o de diseño de las políticas públicas.

Es importante señalar que dentro de este tipo de evaluación en particular Osuna (2013) hace una diferenciación e identifica que la evaluación del diseño de las políticas públicas y el diseño en sí mismo, comparten múltiples elementos. Esto se debe a varias razones: en primer lugar: “la evaluación de diseño tiene como objetivo esencial la mejora del mismo; segundo, la evaluación de diseño, y el propio diseño, forman parte, en realidad, de un mismo proceso que es: el de la planificación del programa o la política de intervención. Una tercera razón estriba en el hecho de que, en el fondo, los parámetros y criterios metodológicos han de ser semejantes siempre que el equipo planificador que diseñe una intervención pretenda que esta sea racional y coherente” (Osuna, 2013: 42).

Es decir, se refiere a que el diseño o la evaluación de diseño, centrará su labor en determinar si efectivamente se cumplen tales exigencias de racionalidad y de coherencia, entre la problemática y las acciones implementadas, como también los objetivos con las metas a alcanzar.

En torno, a un enfoque metodológico se fundamenta con los siguientes autores:

<b>Autor</b>	<b>Enfoque</b>
<i>Rossi (1989)</i>	En el análisis que desarrolla identifica que las relaciones que existe una relación entre la identificación de problemas, el diseño y la planificación de programas y las tareas de evaluación.
<i>Osuna y Márquez (2000)</i>	Cuando hacen referencia a la evaluación conceptual o de diseño, la introducen desde la perspectiva de la planificación. Postulan que, tanto la evaluación, como la planificación, comparten elementos esenciales, tanto que consideran “se convierten en procesos relacionas que caminan al par”.
<i>Reboloso, Fernández y Cantón (2008)</i>	Estudian lo que denominan como evaluación de entrada de forma conjunta con la planificación del programa, aunque consideran que los planificadores, como diseñadores de las políticas públicas, se nutren de los procedimientos formales de la evaluación.

Elaboración propia con base en Cardozo 2006.

En el caso mexicano, este se incorpora de manera sistemática a los procesos de evaluación casi a finalizar el siglo XX. Los tipos de evaluación se diversificaron al reestructurarse el sistema con la creación a nivel federal del Consejo Nacional de Evaluación de Política de Desarrollo Social (CONEVAL) en 2006.

En materia de evaluación, el CONEVAL es la institución responsable de emitir los Lineamientos de Evaluación que deberán observar las dependencias y entidades responsables de operar los programas sociales. Asimismo, tiene a su cargo las evaluaciones de la política y programas de desarrollo social que pueden ser realizadas por la misma entidad, o a través de uno o varios organismos independientes del ejecutor del programa, mediante los cuales podrá emitir recomendaciones respecto al cumplimiento del objetivo social de los programas, metas y acciones de la Política de Desarrollo Social.

En relación con la conceptualización a la Evaluación de diseño, para la institución la “La evaluación de diseño de los programas de desarrollo social es un instrumento dirigido a los programas de nueva creación. Durante el primer año de implementación de un programa debe realizarse este tipo de evaluación. La información que brinda la evaluación de diseño permite tomar decisiones para mejorar la lógica interna de un programa, es decir, saber si su esquema actual contribuye a la solución del problema para el cual fue creado” (CONEVAL: 2018).

Las evaluaciones de diseño se llevan cabo bajo la supervisión de Coneval, siguiendo la metodología de la Matriz de marco Lógico (MML). Estas evaluaciones se elaboran a partir de un cuestionario que busca resaltar la calidad en la definición de tres aspectos: los objetivos, la población objetivo, y la lógica horizontal y vertical del programa (CONEVAL, 2015).

Para el CONEVAL, el objetivo principal de realizar evaluaciones es mejorar políticas públicas y fomentar la rendición de cuentas de quien ejerce la función pública. Por ello ofrece información de los objetivos nacionales y sectoriales de los cuales contribuyen cada uno de los programas y la relación que guardan con otros programas federales que buscan resolver problemáticas afines.

Hay que considerar que cualquier tipo de evaluación, debe generar resultados que tengan efecto directo en el programa evaluado; es decir, los hallazgos que se obtengan permitan la toma de decisiones basadas en evidencia para efectuar cambios a un programa o acción gubernamental.

### **3.3 Características de la Evaluación del diseño del programa de inclusión social “PROSPERA 2016”**

Con el propósito de contribuir al mejoramiento de los programas y las acciones federales de desarrollo social, el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) presento la Evaluación de Diseño del programa de

Inclusión Social PROSPERA 2016, realizada por la instancia evaluadora externa Althiria Consulting.

Al ser una evaluación realizada por equipos evaluadores externos contratados por la instancia, podría pensarse que tiene mayor ventaja por el hecho que el equipo de evaluación externo suele tener un conocimiento más especializado para realizar la evaluación y que actuará con mayor grado de independencia. Empero, a pesar de haber realizado la evaluación una instancia externa, su elaboración se apegó exclusivamente a los Lineamientos generales para la evaluación de los programas federales que establece CONEVAL.

Es importante mencionar que los Lineamientos establecen que los programas sociales deben someterse a evaluación dentro de su primer año de operación, sin embargo PROSPERA al tener 21 años de operación, lo que hace CONEVAL es un análisis de gabinete, que consta de 30 preguntas incluidas en siete apartados: 1) justificación de la creación y del diseño del programa; 2) Contribución a las metas y estrategias nacionales; 3) Población potencial, objetivo y mecanismos de elección; 4) Padrón de beneficiarios y mecanismos de atención; 5) Matriz de indicadores para Resultados (MIR); 6) Presupuesto y rendición de cuentas, y 7) Complementariedades y coincidencias con otros programas federales.

Para el análisis de la Metaevaluación, se comenzará por examinar el objetivo de la evaluación de diseño del programa, continuando con las 30 preguntas de gabinete, se realizará la identificación y valoración de sus criterios empleados para la evaluación de diseño, de igual manera se someterá a juicio su metodología utilizada y por último se valorarán los resultados presentados de dicha evaluación. Es importante recordar que lo anterior mencionado, se realizará con un enfoque sustentando en los criterios planteados por Stufflebeam (validación interna y externa, fiabilidad, objetividad, pertinencia, trascendencia, alcance, credibilidad, oportunidad, accesibilidad, costo/efectividad).

Dado que se trata de un trabajo de gabinete, el proceso de la evaluación se redujo al análisis de documentos. Es decir, más allá de una evaluación en el sentido amplio de la palabra, solo se basó en hacer una adecuada revisión del diseño sin que hubiera relevancia, pertinencia, consistencia y congruencia interna de los objetivos.

Por ello es importante mencionar que, aunque puede parecer que existen múltiples motivos para plantear la realización de una evaluación es necesario considerar que:

- La evaluación no es un fin en sí mismo
- La evaluación debe ser un instrumento para la mejora de las políticas Públicas

### **3.3.1 Objetivo**

La evaluación de diseño del programa de inclusión social PROSPERA 2016, plantea “el objetivo de proveer a la Coordinación Nacional PROSPERA Programa de inclusión Social información que permita mejorar su lógica interna, gestión y resultados”.

A pesar que la evaluación estableció un objetivo para hacer recomendaciones respecto al programa; al analizar la evaluación se encontraron deficiencias en las que con base en los siguientes criterios:

- 1) Pertinencia y credibilidad: se identifica contradicción al interior del texto entre los objetivos establecidos y los medios para lograrlo, y
- 2) Validación Interna y Externa, Fiabilidad y Credibilidad: falta de sustento analítico-teórico en la toma de decisiones para el planteamiento de los objetivos de la evaluación.

Stufflebeam (1974) plantea que una buena evaluación requiere que sus propios logros sean evaluados, ya que muchas cosas pueden ser y frecuentemente son hechas mal en el trabajo evaluativo. Tal es el caso de la evaluación analizada,

ya que, al no formular objetivos claros, da pauta a que se dificulte el control y la evaluación de los resultados, de esta manera no se puede comprobar el nivel que se alcanzó en función de las metas.

Es decir, al solo incluir un objetivo y un solo componente, fue muy sencillo que la evaluación, en términos generales, cubriera todos los aspectos requeridos, aunque en algunos casos se limitara a señalar que esta información no está disponible. Se percibe una tendencia a afirmar el cumplimiento de cada punto, aunque en su fundamento esto quede parcialmente cuestionado.

En términos de Chemlisky (1998), todos los programas son evaluables y en ausencia de objetivos claros en la teoría, el evaluador debe construirla. Es por ello que no se ha podido realizar una evaluación idónea ni correcta en términos generales; pues buena parte de la fundamentación realizada por la evaluación consiste en la reproducción de textos legales, planes y programas relativos al tema puesto que no se consultaron documentos no institucionales.

### **3.3.2 Criterios**

La construcción de criterios es un proceso basado en la teoría, ya que de esa manera se pueden ligar los conceptos a la evidencia empírica. Contrario a lo anterior el CONEVAL señala:

...Cada tema consta de preguntas específicas, las cuales deberán responderse mediante un esquema binario (Sí o no). Para determinar esta respuesta el evaluador deberá fundamentarla con evidencia documental y análisis de la misma, que sustente y justifique los principales argumentos de cada una de las preguntas y temas por analizar...

Esto podría interpretarse como la reducción de la realidad social a una dicotomía y de la evaluación científizada<sup>6</sup>.

La evaluación de diseño que se está enjuiciando se establecieron cuatro criterios (cobertura, focalización, planeación estratégica y pertinencia); los dos primeros fueron planificados para los lineamientos de la evaluación de diseño. Empero, estos criterios solo estuvieron orientados para la cuantificación de la población potencial y objetivo, esto desde el fundamento de los criterios de objetividad, alcance y costo/efectividad; si bien, cumple con algunos criterios, pero deja de lado a otros que al igual tienen la misma importancia.

Por otro lado, los criterios restantes, no estuvieron proyectados en los lineamientos para la realización de la evaluación, sin embargo, se rescataron como producto del análisis, es decir están orientados para identificar los mecanismos y herramientas de planeación; de establecimiento y la definición de metas e indicadores, dando pauta a medir rigurosamente su impacto, la difusión y utilización de los resultados y evaluaciones; privilegiando la coherencia entre la problemática y las acciones.

Sin embargo, los criterios solo funcionaron para los evaluadores como un requisito que CONEVAL establece; teniendo como consecuencia las siguientes inconsistencias:

- 1) con base en las 30 interrogantes con las que se realizó la evaluación de diseño existe incumplimiento en los criterios establecidos;
- 2) Los resultados manifestados en el informe por parte de los evaluadores, no son el fruto del haber enjuiciado al programa, más bien, es el efecto de una síntesis analítica a partir de la realización de las 30 interrogantes aplicadas.

---

<sup>6</sup> Se utiliza este término con la intención de diferenciar un proceso que pretende ser científico, a pesar de no contar con una teoría que la respalde, y el uso de herramientas como sinónimo de metodología.

Una vez identificadas las deficiencias, se puede afirmar que la evaluación que se llevó a cabo es exclusivamente para cumplir con lo establecido en la Ley General de Desarrollo Social, en sus artículos 72 al 80; pues establece que el objetivo de las evaluaciones es revisar periódicamente el cumplimiento de los programas. Apegándose a la vez a los Modelos de Términos de referencia sin tener pauta a implementar los que sean convenientes para la misma.

Es importante mencionar que para llevar a cabo una evaluación correcta que cumpla con el objetivo de identificar hallazgos y así generar recomendaciones a partir del análisis de la congruencia del diseño del programa, debe apegarse a un sustento teórico, en este caso, Stufflebeam implementa once criterios que generaran (validez interna y externa, la confiabilidad, la objetividad, la relevancia, la amplitud, la importancia, la oportunidad, la vigencia prolongada y el costo/ efectividad). La aplicación de esta serie de criterios conforma la estructura esencial para los evaluadores, lo cual permite ampliar el contexto evaluativo, se valora de una manera más completa; por ende, los resultados son más pertinentes y el margen de error disminuye.

Es evidente y fundamental basar la construcción de los criterios en marcos teórico-referenciales que sirvan para la realización de evaluaciones idóneas.

### **3.3.3 Metodología**

El CONEVAL emplea como supuesta metodología, lo que el artículo 3 fracción VII de los Lineamientos Generales para la Evaluación de los Programas Federales de la Administración Pública Federal define como marco lógico, en cual consiste en realizar una matriz de indicadores mediante la cual se describe, en fin, propósito, componentes y actividades. Adicionalmente diseña los indicadores y define las metas, medios de verificación y supuestos para cada uno de los objetivos. Metodológicamente, lo anterior puede ser seriamente cuestionado. En primer lugar, pareciera que la metodología sugerida para realizar las evaluaciones se encuentra

dentro del paradigma cuantitativo. Esto contrasta con la finalidad de los programas sociales.

A pesar que CONEVAL afirma que uno de los principales retos en materia de evaluación es identificar hallazgos con metodologías rigurosas, para determinar si los programas sociales cumplen con el propósito para los cuales fueron diseñados; la metodología empleada en el desarrollo de la evaluación se caracterizó por no tener una técnica de análisis ni de método. Dicho esto, es importante mencionar que, en los términos de Referencia para la Evaluación de diseño, establecidos por CONEVAL (2016) no hacen referencia a métodos sino a fuentes de información y técnicas de investigación y recopilación de información, lo cual se traslada al informe final que presentan los evaluadores.

Es importante mencionar que cualquier tipo de evaluación debe caracterizarse por una metodología robusta y para ello necesita la disponibilidad de recursos para llevarla a cabo, tanto financieros como humanos, técnicos y de información. Pues bien, el tratamiento limitado de la metodología puede suponer una merma de la calidad y orientación evaluativa, aunque en esta evaluación no se haya observado una incidencia negativa significativa conlleva a dificultar la comprensión de los lectores a cuanto a poder calibrar los resultados en función del método aplicado.

En relación con la información obtenida del informe, el apartado dedicado a la metodología fue establecida apegándose una vez más a los modelos de términos de referencia, sin embargo, el énfasis se pone en el objeto de la evaluación más que en una descripción del enfoque metodológico adoptado, teniendo así, deficiencias que a su vez generan que la evaluación no tenga claridad, calidad, ni credibilidad, para fundamentar ni respaldar el análisis realizado y por ende las aportaciones hechas al programa evaluado.

La importancia de la metodología; radica desde la utilización de la misma, pues no debe limitarse a un solo método, sino que existe una variabilidad, para la

consideración del método más adecuado, el cual será fijado por la necesidad y los elementos que lo sustenten; si bien, la metodología permite caracterizar los elementos más importantes, dando pauta a una descripción que sirva como fundamento para el análisis de la evaluación; a través de ella, se puede focalizar la orientación del resultado proyectado.

### **3.3.4 Contenido / desarrollo**

En el informe de la evaluación se desglosó en siete apartados para dar respuesta a las 30 preguntas ya mencionadas para el análisis de gabinete.

1. Resumen Ejecutivo
2. Introducción
3. Descripción General del Programa
4. Evaluación
5. Valoración final del programa
6. Principales fortalezas, Oportunidades, Debilidades, Amenazas y Recomendaciones
7. Conclusiones

A pesar de llevar a cabo un proceso organizado que muestra el informe, se trata de un trabajo de baja calidad por la información ofrecida a lo largo del desarrollo. El equipo evaluador ofrece respuesta a cada apartado dando un diagnóstico excelente con pocas deficiencias, reiterando lo mismo en varias de ellas, lo que nos permite afirmar su poca credibilidad a la hora de evaluar.

Se debe tener en cuenta que el plantear una evaluación requiere preguntarse si este tipo de proceso va a proporcionar la información que se necesita en un momento concreto. Por ello resulta clave entender la utilidad que va a tener la evaluación. En el caso de la evaluación realizada no se tomó en cuenta el propósito de tener una evaluación que facilitara la toma de decisiones.

El contenido de la evaluación fue insuficiente, dado que no incluyó un análisis conceptual, no conllevó un estudio de la racionalidad ni de coherencia. Dando como resultado que al analizar la calidad del diagnóstico para establecer las recomendaciones adecuadas no tuviera adecuación entre el objetivo que se quería lograr y como se llevó a cabo.

Al enjuiciar este apartado de la evaluación toma gran relevancia, mencionar que toda evaluación debe contener un conjunto de preguntas de evaluación que funcionan como eje central de la evaluación. En este caso al no tener autonomía los evaluadores y ajustarse solo a los lineamientos de CONEVAL, generó inconsistencias en cada uno de los apartados que pretendieron llevar a cabo.

### **3.3.5 Conclusiones Generales**

Las conclusiones que presenta el informe se caracterizan más por un hacer diagnóstico del problema. A pesar que el desarrollo se caracteriza por tener deficiencias, las conclusiones que hace la instancia evaluadora son pertinentes y con gran relevancia para fortalecer el programa.

Conclusiones Generales de la evaluación:

- No se cuenta con un documento de planeación estratégica donde se establezca cuál va a ser la evolución del Programa, se especifiquen las metas a corto, mediano y largo plazo, y su contribución al desarrollo social del país.
- No se cuenta con evidencia empírica que demuestre que el Programa ha permitido a la población beneficiaria movilidad social hacia estratos económicos superiores.
- No se cuenta con una evaluación integral de los procesos del Programa.
- No se cuenta con una evaluación de calidad del Programa, donde se establezca el nivel de satisfacción de los beneficiarios y las áreas de oportunidad que ellos consideren se deben instrumentar.

Dicho esto, podemos sostener, basándonos en Scriven (2001) al establecer los criterios en la Key evaluation Checklist, que la evaluación llevada a cabo presentó deficiencias en su diseño y conceptualización relacionadas con su racionalidad, pertinencia y coherencia, así como en su implementación, derivadas en este caso de desajustes entre lo previsto y lo realizado. Tanto una como otras inciden en su evaluabilidad y en los encargos de evaluación.

Si bien, actualmente existen altas expectativas sobre la capacidad que tiene la evaluación a través de sus recomendaciones que retroalimenten y mejoren los programas evaluados. Empero, a falta de evidencia empírica y pertinencia que las sustente no tienen la aportación adecuada.

En el caso de la Evaluación en comento es importante mencionar que necesita un marco de evaluación comprensivo y completo que permita analizar los avances o retrocesos que el programa tiene. Debe establecerse metas claras y exigentes que permitan medir el esfuerzo potencial de la evaluación. Así como indicadores que muestren avances significativos en la consecución de los objetivos de la evaluación.

A pesar que la evaluación no contó con todo lo antes mencionado, las conclusiones generales se caracterizaron en demostrar las deficiencias del programa como tal, sus insuficiencias e irregularidades que la creación del programa tiene. Lo que contribuye afirmar la incongruencia que existió en los apartados que se analizaron a lo largo de este capítulo.

### **3.4 Conclusiones del análisis de la evaluación de diseño del programa de inclusión social prospera 2016**

Ante la evidencia de que los resultados finales, obtenido de la Evaluación de Diseño evaluada, no fueron uniformes para atender recomendaciones, ni mejorar el programa PROSPERA, se crea y justifica la necesidad de realizar una

Metaevaluación, es decir, una nueva evaluación de dicha evaluación. Justificación que encontró sustento en el hecho de que, al momento de realizarse el análisis de la evaluación, los resultados, así como las recomendaciones mostraron deficiencia; derivado que solo consistió exclusivamente en un análisis de gabinete, es decir, solo sigue un modelo de Términos de referencia que da respuesta a 30 preguntas, incluidas en 7 apartados.

Evidentemente, la evaluación de diseño del Programa de Inclusión Social Prospera 2016, y con base a los criterios de: validación interna y externa, pertinencia y oportunidad, se concluye que, básicamente se fundamentó en información institucional, es decir, se limita a unos pocos indicadores que no son capaces de medir suficientemente la calidad de la creación del programa, no puede ser más que un mero formalismo de CONEVAL plagado de repeticiones, que sólo informa del apoyo que otorga y quienes son los beneficiados, del presupuesto ejercido y de algunas acciones complementarias.

Este resultado es responsabilidad principal de la vaguedad de los lineamientos para caracterizar este tipo de evaluación (diseño). Pues simplemente se basa en saber si su esquema actual contribuye a la solución del problema para el cual fue creado. Por lo tanto, es importante resaltar la importancia y la necesidad de usar criterios que permitan elevar el grado de calidad en las evaluaciones.

Es importante mencionar que el Consejo Nacional de Evaluación (CONEVAL) al no presentar modificaciones en cuanto a tiempos, formatos; y regirse bajo un entorno institucional que presta más atención al cumplimiento de los procesos administrativos, genera que los evaluadores y los operadores de los programas tengan más interés en que las evaluaciones arrojen datos positivos que en obtener lecciones útiles para mejorar el desempeño del programa. Esta crítica es realizada en función de los criterios de credibilidad y costo/efectividad.

Es menester mencionar, que las deficiencias se detectaron en cada apartado desagregado de la evaluación hecha con base en los criterios de fiabilidad y objetividad, son:

- La definición de los objetivos que se persiguieron en la evaluación
- La elección del momento para realizar la evaluación
- La predeterminación de los instrumentos para evaluar en la evaluación
- La pertinencia del sistema de evaluación diseñado con lo que se entiende como patrón de calidad
- El cumplimiento por los evaluadores seleccionados con aquellas aptitudes y actitudes establecidas para ejercer como tal: justeza, objetividad, honestidad, capacidad, etc.
- El nivel de participación de los interesados en el proceso de evaluación.
- Las técnicas empleadas por los evaluadores para la recogida de información
- La calidad del informe final presentado por los evaluadores.

Como recomendaciones de la Metaevaluación llevada a cabo se proponen los siguientes puntos:

1. Pertinencia en los hallazgos, es decir, que estén conectados con la lógica causal de la política pública, y no solo aludan a aspectos secundarios que no sean relevantes.
2. Factibilidad de las recomendaciones, de acuerdo con las capacidades administrativas, legales, organizativas y presupuestales de los usuarios.
3. Calidad y precisiones de las recomendaciones específicas al programa deben ser puntuales para definir no solo los propósitos sino los instrumentos que podrían emplearse.
4. Mayor difusión de resultados
5. Juicios de valor explícitos y fundamentados

6. Cultura de evaluación (incentivar a las dependencias y entidades a utilizar las evaluaciones para mejorar la operación de los programas o el uso de recursos.
7. Cambios en los Modelos de términos de referencia y lineamientos
8. Sustento teórico
9. Pertinencia en la elección de criterios.
10. Aprovechar los recursos humanos especializados de las universidades públicas para la realización de las evaluaciones.
11. Formación de recursos humanos en el campo de la evaluación, que cuenten con elementos teóricos y empíricos.

Lo cierto es que los evaluadores tienen que basarse en la investigación social y estar familiarizados con la literatura de evaluación. Es decir, comprender la teoría de la evaluación que están utilizando. Esto requiere experiencia tanto en el tema central de la evaluación como en la problemática observada en la realidad. Por último, pero no menos importante debe tenerse plena conciencia que los hallazgos que se obtengan de una evaluación, así como de una Metaevaluación deben transformarse en acciones de mejora que contribuyan a mejorar los resultados de un programa.

Es menester, mencionar que el primer problema con el que hay que enfrentarse a la hora de abordar una metaevaluación es el de los criterios conforme a los cuales se evalúen las evaluaciones. Si bien, como hemos abordado en capítulos anteriores aún son escasas las propuestas teóricas formuladas de manera específica para la metaevaluación.

Sin embargo, la teoría presentada, así como los instrumentos presentados para llevar a cabo una correcta Metaevaluación son relativamente sencillos. Una primera idea para que las Evaluaciones sean evaluadas y no solo sean aplicaciones de modelos teóricos, consiste en utilizar de forma selectiva el conjunto de listas de comprobación que derivan del sustento teórico presentado en

la presente tesis. Es decir, corresponde a queda evaluador o metaevaluador reconocer y seleccionar las que mejor le convengan según el tipo de evaluación y su objetivo.

A su vez el metaevaluador que se acerca a su objeto con la voluntad de realizar una evaluación crítica, como lo propone Stufflebeam, puede elaborar su propia lista, reduciéndola con base en el conocimiento del conjunto de criterios que le sean más apropiados para llevar a cabo la correcta evaluación. En segunda idea es que, al aplicar el procedimiento sobre las evaluaciones, deberán revisar su pertinencia, su viabilidad, confiabilidad, objetividad y relevancia como lo sugerimos en el planteamiento del problema.

Si bien, al llevarlas a cabo con las recomendaciones ya mencionadas dará pauta a que las metaevaluaciones ya existentes pasen de ser solo aplicaciones de modelos teóricos para la realización de las evaluaciones, a ser un recurso o alternativa práctica ante posibles discrepancias sugeridas después de realizar una evaluación, es decir, con la metaevaluación no sólo se protegerá el buen uso de información y recomendaciones recogida en la evaluación, sino también se protegerá a los objetos evaluados y a todas aquellas personas que intervienen en los procesos evaluativos, tanto evaluados como evaluadores. Su finalidad principal es el perfeccionamiento de los procesos y procedimientos de la evaluación y en última instancia del objeto evaluado, tomando en consideración que la Metaevaluación genera un círculo autoreferente sobre la calidad tanto del objeto evaluado como del metaevaluado y posibilitan su desarrollo hacia el perfeccionamiento.

Finalmente, desde un punto de vista lógico, la Metaevaluación es el cierre de la evaluación, como actividad finalista. La evaluación sin Metaevaluación es parcial, deficitaria e inconclusa. Este es un efecto del imperativo profesional sugerido y mencionado en la presente tesis por Scriven en (1994). La Metaevaluación es necesaria tanto lógicamente y procedimentalmente, como socialmente.

## BIBLIOGRAFÍA

- Bustelo, M. (2003), “¿Qué tiene de específico la metodología de la evaluación?”, en Bañón (Comp.) *La evaluación de la acción y de las políticas públicas*. Madrid.
- Cardozo, M. (2006). *La evaluación de políticas y programas públicos. El caso de los programas de desarrollo social en México*. Colección Conocer para Decidir. México: Cámara de Diputados/Miguel Ángel Porrúa.
  - (2009). *Evaluación y metaevaluación en los programas mexicanos de desarrollo social*, Colección Teoría y Análisis. México: UAM-Xochimilco.
  - (2011), “Organización y métodos alternativos en la evaluación de políticas y programas sociales. La experiencia consolidada y la emergente en el caso de México”, *Revista Gestión y Análisis de Políticas Públicas*, núm. 5, pp. 43-59.
  - (2012a), *Evaluación y Metaevaluación en las políticas y programas públicos. El estado del arte*. México: UAM-Xochimilco.
  - (2012b), *Sistematización de la experiencia de evaluación de Evalúa DF en 2008. Metaevaluación y autoanálisis*. México: UAM-Xochimilco/Evalúad.
  - (2013), “De la evaluación a la reformulación de políticas públicas”, *Política y Cultura*, núm. 40, México: UAM-Xochimilco, pp. 123-149.

- Centeno, Helen (2002), “Meta-evaluación del Programa Antidrogas y Salud Sexual y Reproductiva para Adolescentes, Jamaica 1999-2002: lecciones aprendidas y contribución metodológica”, tesis del Master Profesional en Población y Salud, Costa Rica: Universidad de Costa Rica.
- Cook, T.D. y Gruder, C.L. (1978). Metaevaluation research. *Evaluation Quaterly*, 2, 5 – 51.
- Díaz de Santos. (2004), *La evaluación de políticas públicas de igualdad de género en España*. Madrid: Catarata.
- Díaz, L.F. (2002), “La Metaevaluación y su método”, *Revista de Ciencias Sociales*, año XLIII, núm. 92/93. Costa Rica: Universidad de Costa Rica.
  - (2005), “Metaevaluación del modelo del Sistema Nacional de Evaluación de Costa Rica”, ponencia presentada en el *X Congreso Internacional del CLAD*. Santiago de Chile.
- Díaz, L. F. y R. Rosales. (2003), *Metaevaluación: evaluación de la evaluación de políticas, programas y proyectos sociales*. San José de Costa Rica: Editorial Universidad Estatal a Distancia.
- Dopico, Ileana (s/f), “Metaevaluación: ¿por qué y para qué?” <http://www.sld.cu/galerias/pdf/sitios/prevemi/metaevaluacion.pdf>, consulta 29 de marzo 2018.
- *Evaluation Thesaurus*. Newbury Park: Sage Publications. (2009), “Meta-evaluation revisited”, *Journal of Multidisciplinary Evaluation*, vol. 6, núm. 11, pp. iii-viii.

- Feinstein, O. (2002), "Use of evaluations and the evaluation of their use", *Evaluation*, vol. 8(4), pp. 433-439.
- García, E. (2009), "Metaevaluación de políticas públicas: una visión desde la ciencia política", *Revista Reforma y Democracia*, núm. 43. Venezuela: CLAD
- García, R. y M. García (2010), *La gestión para resultados en el desarrollo. Avances y desafíos para América Latina y el Caribe*.
- Guba, E. y Lincoln (1989), *Fourth generation evaluation* Newbury Park: Sage Publications.
- Guba, E.G. (1975), "Problems in utilizing the results of evaluation", *Journal of Research and Development in Education*, vol. 8(3), pp. 42-54.
- House, Ernest. R. (1994): *Evaluación, ética y poder*, Madrid, Morata.
- Joint Committee on Standards for Educational Evaluation (1994), *The program evaluation standards: How to assess evaluations of educational programs*. Thousand Oaks, CA: Sage.
- Lacasse, M.F. (1993), "Gestion et émasculatión des évaluations: le cas de Canada", *Revue Française d'Administration Publique*, núm. 66, pp. 274-276.
- Lascoumes, Pierre (1998), "Pratiques et modèles de l'évaluation" y "Usages et retombées de l'évaluation", en Kesller, Marie Christine y otros, *Evaluation des politiques publiques*, París: L'Harmattan.

- Mejía, J. (2003), *La evaluación de la gestión y las políticas públicas*. México: Miguel Ángel Pomúa.
- Monnier, E. (1995), *Evaluación de la acción de los poderes públicos*. España: Instituto de Estudios Fiscales.
- Ospina, Sonia, 2001, “Evaluación de la gestión pública, conceptos y aplicaciones en el caso latinoamericano”, en revista *Reforma y Democracia*, núm. 19, Venezuela, CLAD.
- Osuna, José Luis, 2003, “La evaluación participativa, una visión integral y pluralista”, en Bañón (comp.), *La evaluación de la acción y de las políticas públicas*, Madrid, Díaz de Santos.
- Pacheco, J. F. (2010), “Metaevaluación en sistemas nacionales de inversión pública. Santiago de Chile, CEPAL, Serie Manuales Núm. 64.
- Patton, M.Q. (1978), *Utilization focused evaluation*. Thousand Oaks: Sage Publications.
- Ramos, J.M. (2012), *La creación de capacidades institucionales en materia de evaluación: la experiencia mexicana (2007-2012)*. México: El Coleflapem.  
(2013), *Evaluación del desempeño: gestión y política para la generación de resultados efectivos*. México: Auditoría Superior de la Federación.
- Reboloso, E. (1994). La evaluación del personal profesional: ¿Un modelo de evaluación nuevo o una síntesis de modelos previos? *Revista de Psicología Social*, 2, 231-255.

- Rossi, P.H., Freeman, H.E., (1979), *Evaluation: A Systematic Approach*, Beverly Hills: Sage.
- Sanín Ángel, Héctor (1999), "Control de gestión y evaluación de resultados en la gerencia pública: metaevaluación-mesoevaluación", *Serie Manuales - CEPAL*, N° 3 (LC/L.1242-PLC/IP/L.164), Santiago (CEPAL). Publicaciones de las Naciones Unidas, N° de venta S.99. II. G.25.
- Santos, M.A. (1990). *Hacer visible lo cotidiano. Teoría y práctica de la evaluación cualitativa de centros escolares*. Madrid: Akal.
- Schwandt, T. y E. Halpern (1988), *Linking auditing and metaevaluation*. Newbury Park: Sage Publications.
- Scriven, M. (1967), "The methodology of evaluation", end Tyler *et al*, *Perspective of curriculum evaluation*. Chicago: Rand McNally.
- Scriven, Michael, (1969). "An Introduction to Meta-Evaluation". *Educational Product Report*, n.\*2, 5.
- Sime, Luis, (1998). *Metaevaluación: ir más allá de la evaluación para volver sobre ella*. *Educación*. Vol. VII N.\*14. PAG. 199-216.
- Stufflebeam, D. I. (1981). "Metaevaluation: concepts, standards, and uses". In R. A. Berk (Ed), *Educational Evaluation Methodology: The State of the Art*.
- Stufflebeam, D. y A. Shinkfield (1993), *Evaluación sistemática*. España: Paidós.

- Stufflebeam, D.L. (1974). Meta-evaluation. Paper núm. 3. Occasional Paper Series. Kalamazoo: Western Michigan University Evaluation Center.
- Stufflebeam, D.L. y Shinkfield, A.J., (1987). Evaluación sistemática. Guía teórica y práctica. Barcelona: Paidós-MEC.
- Vargas, G. (2007), *Metaevaluación de los programas de la Secretaría de Agricultura (Sagarpa), dirigida a productos agrícola básicos generales*. México: CEDRSSA.