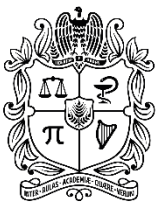


UNIVERSIDAD
NACIONAL
DE COLOMBIA

La efectividad de la evaluación de políticas públicas como uso de la información para la toma de decisiones

Sergio Iván Martínez Porras

Universidad Nacional de Colombia
Facultad de Derecho, Ciencias Políticas y Sociales
Bogotá, Colombia
2019

La efectividad de la evaluación de políticas públicas como uso de la información para la toma de decisiones

Sergio Iván Martínez Porras

Tesis de investigación presentada como requisito parcial para optar al título de:

Doctor en Estudios Políticos y Relaciones Internacionales

Director:

PH.D. André-Noël Roth Deubel

Línea de Investigación:

Estado e Instituciones Políticas

Grupo de Investigación:

Análisis de Políticas Públicas y de la Gestión Pública (APPGP)

Universidad Nacional de Colombia

Facultad de Derecho y Ciencias Políticas y Sociales

Bogotá, Colombia

2019

Dedicada a Dios, nuestra única necesidad, el gozo eternamente renovado, consciente y existente, quien nos brinda la luz que nos fortalece para trabajar por la liberación física, mental y espiritual de la humanidad. ॐ

Resumen

Esta investigación está dirigida a poner a prueba dos de las teorizaciones más importantes de la actualidad acerca de los factores que explican la efectividad de los procesos de evaluación de política pública, vista desde una perspectiva de uso de la información para la toma de decisiones. Desde un marco analítico neoinstitucionalista se contrastan las explicaciones de la OCDE, con el Utilization-Focused Evaluatión de Michael Q. Patton, a partir de datos recolectados en la política de evaluación de políticas públicas en Colombia. Las conclusiones resaltan la importancia de la coordinación entre el proceso de evaluación con el de la intervención evaluada, y una apertura de las organizaciones hacia el aprendizaje y el cambio institucional, por encima de los actuales estándares de calidad promovidos por la OCDE, como la mejor explicación de la efectividad de los procesos de evaluación.

Palabras clave: Uso de la información, toma de decisiones, evaluación de políticas públicas, coordinación, institución.

Abstract

This research is aimed at testing two of the most important theorizations of today about the factors that explain the effectiveness of public policy evaluation processes seen from a perspective of using information for decision-making. From a neo-institutionalist analytical framework, the explanations of the OECD are contrasted with the Utilization-Focused Evaluation of Michael Q. Patton, based on data collected in the policy of evaluation of public policies in Colombia. The conclusions highlight the importance of coordination between the evaluation process with that of the intervention evaluated, and an openness of organizations towards learning and institutional change, above the current quality standards promoted by the OECD, as the best explanation of the effectiveness of the evaluation processes.

Keywords: Use of information, decision making, evaluation of policy, coordination, institution.

Contenido

Pág.

Resumen VII

Introducción 14

1. Literatura y modelo analítico 16
1.1. Marco teórico-metodológico 17
1.1.1. Teorías neoinstitucionalistas en el contexto de los marcos existentes para el análisis de políticas públicas 17
1.1.2. Modelo analítico para el análisis de política pública: Institutional Analysis and Development (IAD) de Elinor Ostrom 26
1.1.3. La evaluación de políticas públicas como disciplina pragmática 44
1.2. Síntesis del proceso analítico 78
1.3. Conclusiones 81

2. Estudio de caso general de la Política Pública de Evaluación de Políticas en Colombia 82
2.1 Análisis institucional 82
2.1.1 Reglas de alcance 82
2.1.2 Reglas de posición y de frontera 88
2.1.3 Reglas de elección 91
2.1.4 Reglas de agregación 95
2.1.5 Reglas de información 96

2.2 Coordinación entre el proceso de evaluación con el de política pública.....98
2.2.1 Involucramiento de los usuarios previstos.....98
2.2.2 Equipo de trabajo conjunto.....100
2.2.3 Reacción a las demandas, preocupaciones e inquietudes de los usuarios previstos para la construcción conjunta de preguntas de evaluación. 102
2.2.4 Trabajo conjunto sobre objetivos prioritarios y usos previstos.....104
2.2.5 Debates metodológicos entre los usuarios previstos y el equipo evaluador 104
2.2.6 Adaptación conjunta del diseño de la evaluación 105
2.2.7 Trabajo conjunto en la interpretación de hallazgos y la construcción de recomendaciones.....108
2.2.8 Planeación conjunta del seguimiento a la implementación de los hallazgos y recomendaciones de la evaluación.....109
2.2.9 Búsqueda proactiva del uso de la evaluación.....110
2.3 Percepción general acerca de la efectividad de los procesos de evaluación111
2.3.1 Percepción general frente a uso de las evaluaciones.....112
2.3.2 Percepción general acerca de factores que apoyan o restringen el uso de las evaluaciones.....113
2.3.3 Implementación de las recomendaciones de las evaluaciones a SINERGIA en los años 2010 y 2013 116
2.4 Conclusiones 118

3. Efectividad de la evaluación de impacto y de diseño al Programa de Conexiones Intradomiciliarias del Ministerio de Vivienda, Ciudad y Territorio 125
3.1 Presentación del Programa de Conexiones Intradomiciliarias.....125
3.2 Caracterización de las situaciones de acción del proceso de evaluación del Programa de Conexiones Intradomiciliarias. 127
3.2.1 Reglas de alcance y resultados potenciales.....127
3.2.2 Reglas de frontera y participantes.....131
3.2.3 Reglas de elección y acciones 132
3.2.4 Reglas de agregación y control.....136
3.2.5 Reglas de posición.....137
3.2.6 Reglas de información.....138
3.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación del Programa de Conexiones Intradomiciliarias.....140
3.4 Análisis de la efectividad de la evaluación del Programa de Conexiones Intradomiciliarias. 145
3.5 Conclusiones 150

4. Efectividad de la evaluación institucional y de resultados del programa Planes Departamentales de Agua del Ministerio de Vivienda, Ciudad y Territorio.....153
4.1 Presentación del programa Planes Departamentales de Agua 153
4.2 Caracterización de las situaciones de acción del proceso de evaluación del Programa de Conexiones Intradomiciliarias. 154
4.2.1 Reglas de alcance y resultados potenciales.....154
4.2.2 Reglas de frontera y participantes.....156
4.2.3 Reglas de elección y acciones 157

4.2.4 Reglas de agregación y control 161
4.2.5 Reglas de posición 161
4.2.6 Reglas de información 162
4.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación del Programa de Planes Departamentales de Agua 164
4.4 Análisis de la efectividad de la evaluación del Programa de Planes Departamentales de Agua. 168
4.5 Conclusiones 173

5. Efectividad de la Evaluación de Operaciones de la Política de Atención Humanitaria a la Población en Situación de Desplazamiento 175
5.1 Presentación de la Política de Atención Humanitaria a la Población en Situación de Desplazamiento 175
5.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Atención Humanitaria 176
5.2.1 Reglas de alcance y resultados potenciales 176
5.2.2 Reglas de frontera y participantes 179
5.2.3 Reglas de elección y acciones 180
5.2.4 Reglas de agregación y control 184
5.2.5 Reglas de posición 185
5.2.6 Reglas de información 186
5.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la Política de Atención Humanitaria 187
5.4 Análisis de la efectividad de la evaluación de la Política de Atención Humanitaria 190
5.5 Conclusiones 197

6. Efectividad de la Evaluación de Operaciones y de Resultados de la Política de Retornos y Reubicaciones a las Víctimas de Desplazamiento Forzado 199
6.1 Presentación de la Política de Retornos y Reubicaciones 199
6.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Retornos y Reubicaciones 201
6.2.1 Reglas de alcance y resultados potenciales 201
6.2.2 Reglas de frontera y participantes 202
6.2.3 Reglas de elección y acciones 203
6.2.4 Reglas de agregación y control 208
6.2.5 Reglas de posición 209
6.2.6 Reglas de información 210
6.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la Política de Retornos y Reubicaciones 210
6.4 Análisis de la efectividad de la evaluación de operaciones y resultados de la Política de Retornos y Reubicaciones 214
6.5 Conclusiones 218

7. Efectividad de la Evaluación Institucional y de Resultados del Programa de Reparación Colectiva. 220
7.1 Presentación del Programa de Reparación Colectiva 220
7.2 Caracterización de las situaciones de acción del proceso de evaluación del Programa de Reparación Colectiva 221
7.2.1 Reglas de alcance y resultados potenciales 221
7.2.2 Reglas de frontera y participantes 223

7.2.3 Reglas de elección y acciones 224
7.2.4 Reglas de agregación y control 229
7.2.5 Reglas de posición 230
7.2.6 Reglas de información 231
7.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación del Programa de Reparación Colectiva 232
7.4 Análisis de la efectividad de la evaluación del Programa de Reparación Colectiva 234
7.5 Conclusiones 242

8. Capítulo 8: Efectividad de la Evaluación Institucional y de Resultados de la Política de Seguridad Alimentaria y Nutricional. 244
8.1 Presentación de la Política de Seguridad Alimentaria y Nutricional (PSAN). 244
8.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Seguridad Alimentaria y Nutricional. 246
8.2.1 Reglas de alcance y resultados potenciales 246
8.2.2 Reglas de frontera y participantes 248
8.2.3 Reglas de elección y acciones 249
8.2.4 Reglas de agregación y control 254
8.2.5 Reglas de posición 255
8.2.6 Reglas de información 257
8.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la Política de Seguridad Alimentaria y Nutricional 258
8.4 Análisis de la efectividad de la evaluación de la Política de Seguridad Alimentaria y Nutricional 262
8.5 Conclusiones 271

9. Capítulo 9: Efectividad de la Evaluación Institucional y de Resultados de las medidas de defensa comercial (salvaguardia y derecho Antidumpling) 273
9.1 Presentación de las medidas de defensa comercial de salvaguardia y derecho antidumping 273
9.2 Caracterización de las situaciones de acción del proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping. 275
9.2.1 Reglas de alcance y resultados potenciales 275
9.2.2 Reglas de frontera y participantes 277
9.2.3 Reglas de elección y acciones 278
9.2.4 Reglas de agregación y control 284
9.2.5 Reglas de posición 284
9.2.6 Reglas de información 285
9.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping ..286
9.4 Análisis de la efectividad de la evaluación del proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping 290
9.5 Conclusiones 298

10. Efectividad de la Evaluación de Resultados de la Política de Prevención y Control de la Contaminación del Aire 300
10.1 Presentación de la Política de Prevención y Control de la Contaminación del Aire (PPCCA) 300

10.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Prevención y Control de la Contaminación del Aire.....301
10.2.1 Reglas de alcance y resultados potenciales.....301
10.2.2 Reglas de frontera y participantes.....303
10.2.3 Reglas de elección y acciones.....304
10.2.4 Reglas de agregación y control.....309
10.2.5 Reglas de posición.....309
10.2.6 Reglas de información.....310
10.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la PPCCA.....311
10.4 Análisis de la efectividad de la evaluación del proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping.....314
10.5 Conclusiones.....321

11. Análisis comparado.....323
11.1 Valoración de la efectividad de la política pública de evaluación de políticas 324
11.2 Valoración del grado de calidad de los procesos de evaluación.....325
11.3 Valoración del grado de coordinación como patrones de interacción requeridos en los procesos de evaluación realizados por la DSEPP.....334
11.4 Conclusiones del análisis comparado.....335

12. Conclusión.....354

Bibliografía.....361

Introducción

La presente investigación está orientada a determinar en qué medida la política pública de evaluación de políticas en Colombia permite confirmar o falsear la teorización actual acerca de la incidencia de la satisfacción de estándares de calidad, y de la coordinación entre el proceso de evaluación y de política pública, en la efectividad de los procesos de evaluación, vista esta última como el uso de la información en la toma de decisiones de política pública. Esto como un modo de avanzar en la discusión acerca de los factores causales que determinan dicha efectividad.

Para esto es necesario, entre otras cosas, definir si la política refleja de manera satisfactoria esta teorización en su implementación, o, por el contrario, no puede ser considerada como un caso válido de manifestación de la teoría. Así mismo, se parte de poner a prueba la hipótesis de que la satisfacción de los estándares profesionales de calidad en los procesos de evaluación y la coordinación entre los procesos de evaluación y de las políticas públicas evaluadas generan la efectividad de la política pública de evaluación de políticas en Colombia. Para alcanzar el objetivo arriba mencionado, y abordar esta hipótesis, se consideran necesarias las siguientes acciones a modo de objetivos específicos:

1. Describir a profundidad el grado de fallas de implementación que tienen los diversos casos de procesos de evaluación de política pública a comparar.
2. Describir a profundidad el grado de efectividad (como uso de la información) que tienen los casos de procesos de evaluación de política pública a comparar.
3. Brindar una estimación acerca del grado de efectividad actual que tiene la política pública de evaluación de políticas públicas a partir de los casos seleccionados.
4. Analizar la importancia relativa de los actuales estándares profesionales de calidad en materia de evaluación, para la generación de efectividad por parte de los procesos de evaluación de política pública.
5. Analizar la importancia relativa que tienen los componentes de la coordinación entre el proceso de evaluación y de la política pública evaluada para la generación de efectividad por parte de los procesos de evaluación de política pública.

En el primer capítulo se hace una presentación de los antecedentes investigativos, el marco teórico y el modelo metodológico a partir del cual se trabajó la investigación. Se utiliza como marco general de análisis en *Institutional Analysis and Development* (IAD) de Elinor Ostrom, y se utiliza como modelos auxiliares de análisis institucional los estándares de calidad de la OCDE para procesos de evaluación de política, y los lineamientos que brinda el modelo *Utilization-Focused Evaluation* (U-FE) de Patton (2012) acerca de la coordinación como adaptación mutua entre el proceso de evaluación con el proceso de política pública evaluada.

En cuanto a la metodología, se busca complementar un estudio de caso general de la política (que ayude a aclarar el marco institucional), con un análisis comparado entre procesos particulares de evaluación dentro de la política (8 casos de procesos de evaluación específicos). En consecuencia, el segundo capítulo comprenderá la presentación del caso general de la institucionalidad de la política pública de evaluación de políticas, que servirá a modo de verificación de la posibilidad y pertinencia de estudiar las variables seleccionadas en la política implementada en Colombia, así como un marco de referencia institucional para los demás capítulos.

En los capítulos en donde se presentan los casos de procesos de evaluación específicos, se medirán las variables seleccionadas (que operacionalizan las dos variables explicativas), describiendo a profundidad la correspondencia entre el diseño y la implementación de los procesos de evaluación, abordando así el primer objetivo específico. Adicionalmente, en estos capítulos también se describirá a profundidad el grado de efectividad de cada proceso de evaluación, por lo que también tendrán la utilidad de servir al segundo objetivo específico de esta investigación.

Los dos primeros objetivos específicos serán ser abordados al interior de cada uno de los capítulos correspondientes a los casos a comparar. Los siguientes tres objetivos serán tratados en los capítulos finales, en donde se presentará una síntesis de todas las mediciones para tener en cuenta en los análisis comparados, y se muestran sus resultados en un capítulo final de conclusiones.

Dentro de las conclusiones que surgieron en este estudio se encuentra la refutación al supuesto racionalista existente en la literatura de metaevaluación y de factores que determinan el uso de la evaluación, en el cual se considera que los actores sociales y políticos, al pretender maximizar sus objetivos, utilizarán la información que dispongan dentro de un marco institucional determinado, razón por la cual, lo único que se necesita para que la evaluación se use, es garantizar el rigor técnico y científico de la información. En contraposición, esta investigación resalta la importancia de considerar los resultados de las evaluaciones como propuestas de política, que requieren de coordinación, defensa y apoyo continuo por parte de sus creadores. Se concluye una refutación parcial de la hipótesis de investigación. Y finalmente, se brindan algunas recomendaciones para Colombia.

1. Literatura y modelo analítico

Esta investigación tendrá un enfoque centrado en los factores (factor-centric design), que se concentra en la capacidad de explicación de los factores o variables causales de la investigación. Se busca, entonces, determinar el efecto causal de las variables independientes en la variable dependiente[1]. Para esto, y a modo de hipótesis central, se busca poner a prueba si la satisfacción de los estándares profesionales de calidad en los procesos de evaluación y la coordinación entre los procesos de evaluación y de las políticas públicas evaluadas generan la efectividad de la política pública de evaluación de políticas en Colombia.

Poner esta hipótesis a prueba desde un enfoque centrado en los factores, significa que el objetivo general de esta investigación es determinar en qué medida la política pública de evaluación de políticas en Colombia permite confirmar o falsear la teorización actual acerca de la incidencia que tienen la satisfacción de estándares de calidad y, de la coordinación entre el proceso de evaluación y de política pública en la efectividad de los procesos evaluativos. Esta última vista desde el uso de la información como modo de racionalización en la toma de decisiones de política pública.

Teniendo este propósito en mente, es necesario presentar la herramienta teórica de análisis con la cual se pretende alcanzar. El objetivo de este capítulo es determinar el modelo analítico (entendido como la conjunción entre el marco teórico y el metodológico) utilizado para la tesis doctoral, y de manera concomitante, ubicarlo en los debates teóricos acerca del análisis de políticas públicas y, por otro lado, de manera más específica, ubicarlo en el debate acerca de los factores determinantes para que un proceso de evaluación sea efectivo (se utilice su información para la toma de decisiones).

Este capítulo muestra, en primer lugar, la discusión histórica acerca del análisis de políticas públicas, profundizando en la discusión actual entre los distintos neo institucionalismos, así como en la discusión que existe entre estos y los enfoques cognitivo y constructivista/crítico. Una vez realizada esta presentación general sobre las discusiones existentes en materia de análisis de políticas públicas, se continúa con una caracterización a profundidad de los preceptos e hipótesis de los modelos seleccionados (el Institutional Analysis and Development de Ostrom y el Utilization-

[1] Tal como lo describen Gschwend and Shimmelfenning: “We call a research-design factor-centric if one is primarily interested in the explanatory power of causal factors. The goal is to estimate the direction and size of particular causal effect of one or a few independent variables, X_i (i=1,… , n), on a dependent variable, Y, and to assess their robustness (…) Typical research questions of factor-centric designs are: does X_i cause Y or what effect does a X_i have on Y and how much?” (2007, p. 8).

Focused Evaluation de Patton), y se especifican las razones por las cuales se seleccionaron estos enfoques para abordar las variables y los objetivos de investigación.

Posteriormente, se presenta el desarrollo de las discusiones epistemológicas acerca de la evaluación de políticas públicas, con la idea de caracterizar los debates y las distintas posturas que se han constituido alrededor de su función y efectividad, observando la evolución entre los paradigmas positivista, neopositivista y constructivista. A continuación, se aclara la discusión teórica actual acerca de los factores que explican el uso de la información de seguimiento y evaluación, y luego se aclaran las variables normativas y estándares de calidad específicos que se tendrán en cuenta para el análisis, teniendo particular atención al mejor modo en el que esto se acopla dentro de las variables del marco general de análisis de política pública seleccionado. Finalmente se hace una presentación de la metodología de la investigación.

Teniendo en consideración esta estructura, las preguntas que guían la realización de este capítulo son: ¿cuáles son los principales debates epistemológicos acerca de la evaluación de política pública? ¿cuáles son los principales enfoques de análisis para políticas públicas? ¿Cuáles son sus debates? ¿qué modelo analítico va a ser utilizado para la presente tesis doctoral? ¿cuáles son sus ventajas frente a los demás modelos y enfoques de análisis? y ¿Cuál es la metodología que se va a utilizar para la presente investigación?

1.1. Marco teórico-metodológico

1.1.1. Teorías neoinstitucionalistas para el análisis de políticas públicas

El análisis científico de las políticas públicas (en contraposición al arte o la filosofía de gobernar previo a este) tiene como texto fundador *La orientación hacia las políticas* de Harold D. Lasswell (1951). Este autor parte de situarse en un contexto de crisis de seguridad nacional (guerra fría), el cual hace surgir la necesidad de determinar el modo en el que se puede utilizar, lo más eficientemente posible, la fuerza de trabajo y los recursos de EEUU. Plantea que para lograr esto es necesario tomar al proceso de política pública como objeto de estudio, y la racionalización de las decisiones en su interior como el objetivo esencial a obtener.

Lasswell describe que se ha venido desarrollando una orientación hacia las políticas, la cual tiene una doble dimensión: la primera, desarrollar la ciencia de la formación y ejecución de las políticas y, la segunda, mejorar el contenido concreto de la información y la interpretación disponibles para los hacedores de políticas (lo cual debe ser interdisciplinar). De esta orientación se deriva la aspiración a desarrollar el estudio científico de las políticas, lo cual incluiría:

- Los métodos de investigación del proceso de la política.

- Los resultados de los estudios de las políticas.
- Los descubrimientos de las disciplinas que pueden aportar contribuciones importantes para las necesidades de inteligencia del momento.

Este autor, en correspondencia con el auge de la postura positivista en su momento, enfatizó la importancia que han tenido los métodos cuantitativos en el desarrollo de las ciencias sociales, que se evidencia en el desarrollo que tuvo la psicología a partir del uso de test de inteligencia. Sin embargo, también recalcó la fortaleza y utilidad de los métodos cualitativos, que han permitido la última fase del desarrollo de la psiquiatría a partir de la observación meticulosa, la medición y el registro de datos.

Profundizó en la caracterización democrática de la ciencia de las políticas, al manifestar que se deben esclarecer los propósitos valorativos subyacentes en las políticas. Al defender la tradición norteamericana, planteó que todo el énfasis se pone en el desarrollo del conocimiento que permite la realización más completa de la dignidad humana (este sería el valor fundamental y más importante de las ciencias de políticas de la democracia). Esto significó, en contraposición a la neutralidad epistemológica del paradigma postpositivista, aquiescencia para que el científico acepte o rechace proyectos de investigación en consonancia con sus metas valorativas.

Siguiendo una perspectiva pragmática, Lasswell planteó que este enfoque científico de las políticas de la democracia implicaba reelaborar y evaluar el conocimiento procedente de cualquier fuente, que pareciera tener un peso importante en los problemas de las principales políticas del momento. De tal manera, el intercambio de conocimiento entre científicos de ciencias naturales, científicos de ciencias sociales y de hombre de acción resultó deseable, así como necesario. En este sentido, se consideró necesario avanzar en una agenda de construcción de técnicas para una cooperación interdisciplinar fluida.

Esta postura inicialmente racionalista, humanista, conductista y, en general, hasta cierto punto postpositivista, se va a ir transformando y diversificando a lo largo de las décadas. Luego de Lasswell, el estudio de las políticas públicas se fue fragmentando en los distintos sectores (disciplinas) de gobierno:

> "… a medida que el concepto de políticas públicas se fue introduciendo en el lenguaje académico, su comprensión como acción del Estado en un sector particular, el análisis se ha realizado frecuentemente desde los conceptos y las perspectivas teóricas dominantes en la disciplina correspondiente al sector en cuestión" (Roth, 2016, p. 4).

Fue en la ciencia política que el análisis de políticas públicas encontraría asidero, en consecuencia, entró a participar en los debates internos que en esta se presentaba. Las perspectivas más importantes que abordaron como objeto de análisis a las políticas públicas en los años posteriores al texto de Lasswell, fueron

el institucionalismo (que venía de muchos años atrás), el conductismo y el racionalismo.

El institucionalismo planteaba que la ciencia política consistía en un análisis formal y legal (institucional) del Estado, tenía cuatro características principales según André Roth (2015, pp. 3-6):

- Ley como elemento central de estructuración del gobierno e influir en el comportamiento de los ciudadanos y difundir una lógica de lo adecuado.
- La estructura del sistema político es lo que determina el comportamiento de los individuos.
- Comprensión del sistema político solo desde una perspectiva histórica particular en donde el Estado y la sociedad se influyen mutuamente para producir el sistema político y sus decisiones.
- Perspectiva normativa en donde los hechos y los valores están estrechamente entrelazados (rechazo de la separación pregonada por el positivismo).

Estas características del institucionalismo fueron fuertemente criticadas en la década de 1960 por las posturas conductista y racionalista, cuyo eje central de discusión era su poco carácter científico (Peters, 2003, pp. 20-26). Estos fueron los tres paradigmas o posturas que dominaron la ciencia política hasta 1980 (Roth, 2015, p. 3). El conductismo, por su parte, se apoyó en una perspectiva antinormativa, objetivista, positivista, individualista y sistémica. Planteaba que el comportamiento individual se debía principalmente a valores personales y a normas informales, en contraposición a las normas formales defendidas por el institucionalismo, que llegaron a ser consideradas como “cascarones vacíos” que permitían a los individuos tomar decisiones de acuerdo con sus valores y objetivos (p. 4).

Por otro lado, el racionalismo o utilitarismo llevó aún más lejos los postulados conductistas al señalar principalmente, que los individuos y los actores colectivos efectúan elecciones racionales en busca de maximizar su utilidad material e inmaterial (p. 4). Por lo que el interés se convirtió en la variable explicativa principal tanto del cambio de las políticas, como de sus resultados. Tanto desde el conductismo como desde el racionalismo, las instituciones no presentaban una influencia determinante sobre la formación de las preferencias de los individuos (Knoepfel et al., 2006, p. 100).

Desde estas dos perspectivas, surgen teorías políticas que establecen al Estado como una variable dependiente de la sociedad (centradas en la sociedad), entre estas se destacan el pluralismo, el elitismo y el marxismo. En estas, son las fuerzas de la sociedad las que determinan las decisiones y el curso de las acciones públicas; y la administración pública y burocracias son instrumentos neutrales y de ejecución que se encuentran al servicio del gobernante (quien toma las decisiones), más no ejercen influencia sobre el contenido de las decisiones. Por esta razón, la reflexión

acerca de las políticas pública va a centrarse sobre la toma de decisiones (Roth, 2015, pp. 4-5).

En contraposición a esta perspectiva que observa al Estado como dependiente de la sociedad, se desarrolló una perspectiva que postulaba a la sociedad como variable dependiente del Estado (centrada en el Estado), en donde se destacan teorías políticas como el neo-marxismo, el neo-corporativismo, el profesionalismo y el tecnocratismo (Parsons, 2007, p. 206). En esta postura se considera al Estado como una institución autónoma de la sociedad, capaz de incidir en el contenido de las decisiones, que funciona como un selector de demandas sociales y proveedor de servicios (Roth, 2015, p. 5).

A finales de los años 70' y principios de los 80' surge el neoinstitucionalismo como una reacción al conductismo y al racionalismo, el cual intenta tomar en consideración elementos provenientes tanto de la perspectiva centrada en la sociedad, como de la perspectiva centrada en el estado en sus explicaciones acerca del cambio de las políticas públicas y del proceso político (p. 5). Sin embargo, el neoinstitucionalismo no es una perspectiva homogénea, en su interior, se encuentran en discusión las perspectivas económica, sociológica, histórica y discursiva (esta última siendo la más reciente, se puede ver en Boussaguet et al. 2010, pp. 350-359, y Schmidt, 2008). Por esta razón, el neoinstitucionalismo se puede describir mejor como una familia de teorías que tienen en común el énfasis en la importancia de las instituciones para la comprensión de los procesos sociales y políticos (Roth, 2015, p. 3).

Estas perspectivas tienen en común que "Las instituciones políticas tienen una influencia sobre los actores y los procesos sociales (y que, en la medida en que estas) definen las reglas del juego, las instituciones impactarán sobre los resultados (outcomes)" (Boussaguet et al., 2010, p. 357). También tienen en común que observan las relaciones entre Estado y sociedad de manera interdependiente, recíproca, multicausal y con causalidad circular. Por lo que -en su conjunto- es una perspectiva renovada acerca del papel de las instituciones formales e informales en el comportamiento de actores colectivos e individuales y del proceso de formación de estas normas (Roth, 2015, p. 14).

En contraposición a estos puntos en común, cada perspectiva privilegia unos u otros elementos explicativos que en la vida cotidiana se encuentran inextricablemente imbricados y en movimiento, por lo que los investigadores, si bien privilegian a una u otra, es común que complementen su modelo con elementos característicos de otras perspectivas neoinstitucionalistas (p. 14). El siguiente cuadro adaptado, nos permite ver las principales diferencias entre estas perspectivas:

Diferencias principales entre las perspectivas neoinstitucionalistas

Perspectiva	Económica	Histórica	Sociológica	Discursiva
Autores principales	Williamson (1985); Ostrom (1990), North (1990)	Evans et al. (1985); Hall (1986), Steinmo et al. (1992), Weaver, Rockman (1993)	March, Olsen (1989); Powell, Di Maggio (1991); Scott, Meyer (1994)	Hall (1993); Roe (1994), Dryzek (1990, 2000); Campbell y Pedersen (2001); Stone (1988, 1989), Schmidt (2008).
Definición analítica de las reglas institucionales	Contratos o arreglos voluntarios que resultan de las interacciones repetidas entre los actores y que estabilizan las expectativas individuales asegurándoles una cierta previsibilidad en cuanto a resultado de la acción colectiva.	Procedimientos (in) formales, rutinas, normas jurídicas que son incrustadas en las estructuras del sistema político, que reflejan las relaciones de poder entre los actores y que predefinen el acceso a la arena de decisión.	Valores culturales, normas sociales, símbolos, ritos, costumbres, etc. que limitan las capacidades cognitivas de los actores, que definen los roles de los miembros de una organización y que aseguran la legitimidad social de las instituciones.	Constructos políticos y sociales que moldean las interacciones discursivas por las cuales los actores redefinen constantemente sus intereses y normas culturales legítimos.
Estatuto epistemológico de las reglas institucionales	Variable dependiente (nivel micro) Individuos determinan instituciones.	Variable independiente y dependiente (nivel meso) instituciones determinan a, y son determinadas por individuos.	Variable independiente (nivel macro). Instituciones determinan a los individuos.	Variable independiente y dependiente (nivel meso). Instituciones determinan a, y son determinadas por individuos.
Creación de las reglas institucionales	Funcional: reglas institucionales creadas para servir los intereses de grupos concernidos.	Contingente: reglas desarrollándose en el seno de las reglas institucionales existentes.	Inmanente: reglas institucionales (re)producidas por individuos, grupos y organizaciones.	Contingente: las interacciones discursivas producen las reglas institucionales, las cuales, a su turno, determinan las dinámicas e interacciones discursivas de los actores.
Cambio institucional	Apunta a restablecer el equilibrio, si las preferencias de los actores se modifican y al reducir los efectos perversos inducidos por las reglas institucionales anteriores.	Las reglas institucionales tienen ante todo un efecto estabilizador, aunque ciertas constelaciones institucionales ofrecen, a cierto momento, oportunidades de cambio.	Las reglas institucionales influyen en la visión del mundo de los actores que adoptan, en el momento del cambio institucional, un nuevo marco de referencia incursiona entre varios posibles.	Las dinámicas discursivas pueden inducir al cambio o al mantenimiento de las reglas institucionales.

Fortalezas de la perspectiva	Claridad y coherencia conceptual, teoría formal y deductiva.	Combinación de varias lógicas de decisión y toma en consideración de las estructuras.	Foco en la sociología de las organizaciones y en las actitudes y roles sociales.	Establecimiento de un puente entre el racionalismo y el constructivismo. Foco sobre la dinámica de las interacciones entre ideas, formas de deliberación públicas y reglas institucionales.
Debilidades de la perspectiva	Capacidad limitada para explicar los (no) cambios institucionales, límites inherentes a los modelos del "public choice", funcionalismo.	Enfoque inductivo aun no consolidado teóricamente, riesgo de determinismo estructural.	Ambigüedad en cuanto a la verdadera naturaleza de las relaciones entre normas sociales e institucionales políticas (formales).	Enfoque con base en estudios de casos. Particularismo no permite la formulación y comprobación de hipótesis. Riesgo de alejarse de los hechos reales.

Adaptado de Roth[2] (2015, p. 15).

El neoinstitucionalismo económico integra el factor institucional como reductor de incertidumbre y como determinante para la expresión de las preferencias de los actores a la teoría de la elección racional (p. 8). Esta perspectiva añade realismo a la teoría microeconómica de elección racional, puesto que, si bien los individuos procuran maximizar su conducta hacia ordenamiento prioritarios estables y consistentes, lo hacen en presencia de límites cognoscitivos, información incompleta y dificultades en el seguimiento y capacidad para hacer cumplir los acuerdos, por lo que las instituciones surgen y persisten cuando confieren mayores beneficios que los costos de transacción[3] al crearlas y mantenerlas (Powell y DiMaggio, 1999, p. 36). En esta postura, el Estado debe encargarse de disminuir los costos de transacción, los cuales se observan como barreras institucionales que generan distorsiones de mercado, impidiendo una competencia perfecta, la cual, se observa como un objetivo político deseable (Roth, 2015, p. 8).

En uno de los textos fundadores del neoinstitucionalismo económico, Douglas North define las instituciones como las reglas de juego de la sociedad, y como las restricciones diseñadas para la interacción humana, lo que les permite estructurar

[2] Roth elaboró este cuadro con base en Knoepfel (2006: 103) para las vertientes económicas, históricas y sociológicas. Y lo completó con la vertiente discursiva, a partir de Schmidt y Crespy en Boussaguet et al. (2010: 350-359).

[3] Los costos de transacción son los costos de negociación, ejecución y cumplimiento (Powell y DiMaggio, 1999, p. 36). Son los recursos necesarios para medir tanto los atributos físicos como legales de los bienes que se están cambiando, los costos de vigilar, patrullar, y hacer cumplir los acuerdos y la incertidumbre que refleja el grado de imperfección en la medición y el cumplimiento de los términos de intercambio" (North, p. 1990, p. 33).

incentivos en el intercambio económico o político (1990, p. 3). Para este autor, el cambio institucional es formado y conducido a partir de las actividades intencionales llevadas a cabo por las organizaciones, las cuales buscan maximizar su riqueza, ingresos o alcanzar de una mejor manera sus objetivos a partir de las oportunidades que son brindadas por la estructura institucional.

En la búsqueda de estos propósitos, las organizaciones continuamente y de manera incremental, alteran la estructura institucional (p. 73). Esto es así, porque las organizaciones pueden maximizar sus actividades dentro del marco institucional o destinando recursos para cambiar las limitaciones institucionales existentes (p. 79). De manera puntual, para North las organizaciones le dan forma al cambio institucional a partir de: la demanda derivada resultante de inversión en conocimiento de todo tipo; la interacción constante entre la actividad económica organizada, el acervo de conocimientos y el marco institucional; y la alteración incremental de las restricciones informales como un subproducto de maximizar las actividades de las organizaciones (p. 78).

El neoinstitucionalismo histórico se centra -principalmente- en la necesidad de aprehender al Estado en una perspectiva de largo plazo y con método comparado, presta particular atención a las asimetrías de poder que derivan de la acción y el desarrollo de las instituciones. Sus investigaciones muestran el desarrollo institucional como el resultado de una “dependencia de sendero” (describe la existencia de movimientos cumulativos que cristalizan los sistemas de acción y las configuraciones institucionales propias a una política pública o sistema particular), por lo que la herencia política e institucional pesa sobre la selección actual y futura de la acción pública. Su mayor problema su escasa pertinencia y practicidad para la predicción o la prescripción (Roth, 2015, p. 12).

Dos de los autores fundamentales en la definición del neoinstitucionalismo histórico son Steinmo y Thelen. Estos autores identificaron cuatro fuentes principales de dinamismo en las instituciones (1992, pp. 16-17): cambios en el contexto socioeconómico o político que producen situaciones en donde instituciones pasadas se vuelven de nuevo relevantes; cambios en el contexto socioeconómico o de equilibrios de poder en donde se generan situaciones en que pasadas instituciones se ponen al servicio de diferentes fines, en donde entran nuevos actores que persiguen sus fines en el marco de instituciones existentes; cambios exógenos que producen cambios en objetivos y estrategias dentro de los marcos institucionales existentes; por último, la cuarta fuente de dinamismo se refiere a situaciones en las que los actores políticos ajustan sus estrategias para acomodar cambios en las instituciones.

Para estos autores es de particular relevancia las luchas de los actores alrededor de las elecciones institucionales. Esto es así, porque de estas elecciones pueden devenir amplios caminos de política pública (p. 22).Esto está en consonancia con la idea de que existentes estructuras institucionales le dan forma a nuevas disputas alrededor de nuevas políticas e instituciones.

El neoinstitucionalismo sociológico establece que la mayoría de las formas y procedimientos son prácticas de origen cultural parecidas a los mitos y ceremonias inventadas en las diversas sociedades (Powell y DiMaggio, 1999, pp. 34-75). Se integran, entonces, las variables culturales en las condiciones de formación y de funcionamiento de las organizaciones. Las culturas en sí son entendidas como instituciones, por lo que surge la necesidad de analizar elementos cognitivos (entre éstos, las instituciones no escritas o informales) que afectan los comportamientos individuales y que determinan la legitimidad de las organizaciones (Roth, 2015, p. 12). Los principales elementos cognitivos a consideran son los valores, los cuales persisten en el tiempo al ser interiorizados por los individuos, estructuran la vida social y política, determinan el comportamiento de los individuos, y forjan las identidades (Roth, 2015, p. 13).

Dentro de esta corriente se encuentran March y Olsen. Estos autores manifiestan que dentro de los procesos de cambio institucional se pueden incluir transformaciones mundanas y acumulativas de la vida cotidiana, así como metamorfosis más raras en los momentos cruciales de la historia como, por ejemplo, cuando se ponen en cuestión o se cambian valores sociales (1997, p. 261). Estos cambios son producidos a partir de un encuentro entre las instituciones o la acción basa en ellas, y los entornos que consisten en otras reglas.

Los autores aclaran que a medida que una institución desarrolla competencia para actuar en un sistema político particular, ese desarrollo viene acompañado de un incremento aun más significativo de la disparidad entre la adaptación de las instituciones y las demandas sociales cambiantes (p. 264). En este sentido, las ideas entorno al comportamiento pertinente cambia poco a poco, lo que produce retrasos en la adaptación de las instituciones a sus entornos. En este sentido, puede decirse que las instituciones dentro del neoinstitucionalismo sociológico cambian de acuerdo con su grado de idoneidad frente a las percepciones y requerimientos de la sociedad.

Por último, el neoinstitucionalismo discursivo, de influencia constructivista y más nuevo que las demás perspectivas, pretende poner en evidencia el papel específico de las ideas y de las interacciones discursivas en la explicación de los procesos políticos y de las políticas públicas. Autonomiza los elementos discursivos para enfatizar su propia fuerza en la construcción de reglas institucionales, focaliza el rol de las ideas en la construcción de intereses y de valores, y da importancia a la persuasión en los procesos de legitimación de las instituciones (Roth, 2015, pp. 13-14). Estos autores diferencian dos tipos de ideas: las cognitivas que son basadas en la experticia y/o los intereses, y las normativas legitimadas por valores culturales (Boussaguet et al. 2010, p. 352).

La anterior es una caracterización de las discusiones recientes y actuales que existen al interior del neoinstitucionalismo. Es importante considerar que esta última perspectiva discursiva es producto de la discusión que las otras perspectivas

neoinstitucionalistas (histórica, económica y sociológica), han sostenido con dos enfoques de análisis para el análisis de políticas públicas de la década de los 90' hasta la actualidad, estos son: el enfoque cognitivo y el enfoque interpretativo/crítico.

El enfoque cognitivo, cuyos principales exponentes son Pierre Muller (con su análisis de referenciales de política pública, 2002) y Paul Sabatier (con su Advocacy Coalition Framework, 2007), influenciados por los epistemólogos Thomas Khun e Irme Lakatos, ponen resaltan las ideas y valores como elementos a abordar como datos para explicar el cambio en las políticas públicas. Esta perspectiva no rompe con el postpositivismo dominante en el campo del análisis de políticas públicas, sin embargo, va a venir a incluir variables explicativas que hasta la fecha se consideraban como subjetivos -tal como son los referenciales, creencias, valores, ideas, etc.-, a pesar de esto, no ha sido capaz de proponer un modelo normativo para la acción política, lo cual limita su utilidad (Roth, 2016, p. 10).

Poco tiempo después del surgimiento del enfoque cognitivo, surgiría otra ruptura epistemológica, esta vez, de carácter más crítica e interpretativa. El llamado giro argumentativo y/o retórico manifiesta que, si las políticas públicas dependen de las ideas, es necesario dar cuenta del proceso de formación de esas ideas, lo cual hace resaltar el contexto en el que estas surgen (p. 10). Fue Giandoménico Majone (1997) quien vendría a resaltar la importancia de dar cuenta de la manera como se imponen unas ideas sobre otras en un contexto de toma de decisiones, según él, por medio de procesos artesanales de argumentación y persuasión. En este sentido, se hizo necesario reconocer el papel de otros atributos humanos, otras formas de conocimiento, emociones, etc. que son por lo general excluidos de los modelos explicativos neopositivistas al ser tratados como elementos por fuera de la racionalidad o sencillamente subjetivos (Roth, 2016, p. 10).

Esta perspectiva, crítica y constructivista, busca romper con el empirismo postpositivista, radicaliza la importancia de las ideas, la subjetividad y los valores, se centra en develar las asimetrías de poder en el proceso de las políticas públicas, y el cambio en estas se explica más por un cambio en los valores y en los sistemas apreciativos de estas que por cambios en variables objetivas. Esta postura rescata las metas valorativas de la democracia sugeridas en un principio por Lasswell (p. 11). En esta línea Frank Fischer ha establecido la importancia de crear instancias de deliberación participativa, transdisciplinares y que incluyan a todos los involucrados y los saberes locales en igualdad de condiciones, así mismo, plantea un nuevo papel para los expertos de políticas, quienes deberían ejercer una función de facilitadores en contraposición a una de traductores (Fischer, 2009). Por este medio, se busca democratizar el proceso de las políticas públicas para sustraerlo a las élites políticas y o tecnocráticas que dominan o controlan el proceso gubernamental y mediático (Roth, 2016, p. 12).

1.1.2. Modelo analítico para el análisis de política pública: Institutional Analysis and Development (IAD) de Elinor Ostrom

Como ya se manifestó, el objetivo de esta investigación es determinar en qué medida la política pública de evaluación de políticas en Colombia, permite confirmar o falsear la teorización actual acerca de la incidencia de la satisfacción de estándares de calidad y de la coordinación del proceso de evaluación y de política pública, en el uso de la información como modo de racionalización, en la toma de decisiones de política pública (es decir, en la efectividad tanto de los procesos de evaluación como de la política pública evaluada, tal como se desprende de sus objetivos).

El abordaje teórico (compuesto por elementos conceptuales, hipotéticos y deductivos) con el cuál se piensa trabajar este objetivo, parte de reconocer que en gran parte de la literatura en la que se teoriza acerca de la evaluación de políticas públicas, por ejemplo: Kusek y Rist (2005); Morra y Rist (2009); Patton (2002); Weiss (1975); entre muchos otros, por lo general, son textos que brindan lineamientos metodológicos y administrativos para que la evaluación pueda ser usada para la racionalización de decisiones.

Esta tendencia también la siguen los textos acerca de evaluación de diversas organizaciones de cooperación multilateral, e incluso, de organizaciones encargadas de la evaluación de políticas públicas como el DNP en Colombia. A pesar del considerable número de textos que se pueden encontrar sobre lineamientos para una evaluación efectiva, la subdisciplina ha alcanzado un relativo consenso acerca de los principales lineamientos a seguir. Estos han quedado plasmados en listas relativamente homogéneas de “Estándares profesionales de calidad”.

Poner a prueba la teorización acerca del papel de estos estándares de calidad, es la razón de la selección de su satisfacción como primera variable independiente de la hipótesis del presente estudio. En este sentido, es importante aclarar que la relevancia teórica de la presente investigación se centra en evaluar premisas no cuestionadas en gran parte de la bibliografía (King, Verba y Keohane, 2000, p. 27). Sin embargo, se deberá prestar atención a la medida en la cual la política pública de evaluación refleja, de manera satisfactoria, la implementación de la teorización acerca de los estándares de calidad.

Teniendo en consideración que la naturaleza de las variables seleccionadas se relacionan con la obtención de resultados, la incidencia de instituciones particulares y la interacción-coordinación de actores, se retoma, en primera instancia, los postulados pertenecientes al Institutional Analysis and Development (IAD) de Elinor Ostrom, el cual servirá como marco general de análisis de política pública, y que ha sido definido como un marco de trabajo, o una teoría de teorías, que cuenta con

distintos niveles y distintos grados de especificidad teórica que permiten la integración de una familia de teorías[4] (Ostrom, 2015, p. 7).

El marco IAD de Ostrom nos ayuda a direccionar preguntas acerca del cambio (reformas o transiciones) institucionales[5], y a predecir probables resultados de las políticas públicas (Ostrom, 2007, p. 26). Esta segunda utilidad es fundamental para la determinación de la explicación de la efectividad de la política en cuestión, por lo que es, en términos generales, el marco analítico más amplio en el sentido de fenómenos sociales a incluir. Así es propuesto por Ostrom:

> "Para los políticos y los académicos interesados en las cuestiones relacionadas con la forma en la que los diversos sistemas de gobierno permiten a los individuos resolver problemas democráticamente modificando reglas a diversos niveles, el marco ADI ayuda a organizar las capacidades de diagnóstico, análisis y prescripción" (2015, p. 69).

Este marco de análisis es considerado por muchos autores (cómo Roth y Knoepfel tal como se vio en la sección anterior), como perteneciente única o principalmente a la vertiente económica del neoinstitucionalismo, sin embargo, esto es discutible, pues a pesar de incluir elementos matizados de la teoría de la elección racional (especialmente en el modelo del individuo y en la situación de acción), retoma fuertes elementos del neoinstitucionalismo sociológico, histórico, e inclusive del enfoque cognitivo (afirmación que se precisará más adelante).

El neoinstitucionalismo económico ha sido fuertemente criticado por favorecer una visión tecnocrática e inclusive antidemocrática de las políticas públicas, tal como lo plantea Roth citando a Garson, desarrollando su planteamiento acerca de la obsesión metodológica encarnada por la teoría Public Choice a partir de la década de los setenta:

> "… la dominación de la teoría económica en la explicación de los comportamientos individuales y colectivos se expandió a gran parte de las ciencias sociales, y también impuso su punto de vista normativo para la formación de las políticas. Como señala Garson (1992:155), las ciencias de políticas pasaron de querer ser un campo de investigación al servicio de la democracia a ser realmente un oficio al servicio de los gobiernos de turno; un oficio que se realizaba desde una actitud autoritaria –"tiránica"- de imposición de saberes tecnocráticos." (2016, p. 6).

---

[4] Dentro de las teorías que componen el IAD se encuentran: teoría microeconómica, teoría de juegos, teoría del costo de transacción, la teoría de la elección social, teoría de la elección pública, teoría constitucional, teoría de bienes públicos, y teoría de bienes comunes (Ostrom, 2015).

[5] Siguiendo a Ostrom, las instituciones "… son prescripciones que los seres humanos usamos para organizar todas las formas de interacciones repetidas y estructuradas, incluyendo las que acontecen en familias, barrios, mercados, empresas, clubes deportivos, iglesias, asociaciones privadas y gobiernos a todas las escalas" (2015, p. 39).

En este punto, se presenta una primera ruptura con los planteamientos de Ostrom, puesto que su marco de análisis establece claramente la necesidad de democratizar la gestión pública, principalmente por la necesaria interiorización que los individuos deben hacer de las instituciones para que las cumplan, tal como manifiesta acerca de la legitimidad y el cumplimiento de las instituciones:

> "Si los gobernantes imponen las reglas sobre todo mediante la fuerza y el decreto, es improbable que los individuos sometidos a dichas reglas desarrollen parámetros delta[6] internos asociados con su incumplimiento" (2015, p. 231).

Otras razones importantes por las cuales se puede decir que Ostrom defiende la democratización del proceso de política pública son: el necesario sentimiento de los participantes de que las reglas que gobiernan la situación son apropiadas, por el elevado costo de imponer cumplimiento en caso contrario (p. 61); necesaria comunicación constante entre participantes sin la cual los recursos comunes tienden a ser sobreexplotados (p. 147); y la explicación acerca de los comportamientos cooperativos (fundamentales para la acción colectiva):

> "… para que normas (…) lleven al individuo a comportarse como un cooperador condicional y a generar resultados positivos, se necesita algún tipo de acuerdo que agrupe a los cooperadores condicionales (…) un requisito clave para cooperar con éxito es que los cooperadores condicionales sean capaces de encontrarse unos a otros, ya sea mediante vínculos espaciales o institucionales. Si alrededor de los cooperadores condicionales hay demasiados egoístas racionales, la cooperación pueden desmoronarse sin más" (pp. 180-181).

Otra ruptura importante que debe identificarse entre el IAD de Ostrom y el neoinstitucionalismo económico de los años 70' y 80', es la matización que Ostrom realiza a la Teoría de la Elección Racional y a la Teoría de Juegos. Frente a esta última plantea que, si bien es una teoría fuerte y esclarecedora, solo puede elucidar la estructura de situaciones relativamente simples, no de juegos complejos y complejamente estructurados –como por ejemplo el ajedrez- (p. 75).

Lo anterior se explica porque la Teoría de Elección Racional y sus modelos han probado servir para predecir el comportamiento humano en escenarios estables de mercados competitivos y en escenarios competitivos electorales y parlamentarios, en los cuales los individuos pueden aprender la estructura completa de la situación

[6] Parámetros delta son las recompensas o costos relacionados con el cumplimiento o incumplimiento de la prescripción compartida. Los internos son aclarados del siguiente modo: "Se puede pensar en los parámetros delta que se originan en fuentes internas como la culpa o la vergüenza que se siente cuando se incumple una prescripción y el orgullo o "satisfacción" que se siente cuando se cumple, sobre todo si es difícil lograrlo (p. 207).

y vincular las posibles preferencias positivas y negativas a las acciones y los resultados (p. 151). Sin embargo, a medida que el escenario se complejiza, estas dos últimas condiciones para predecir el comportamiento se van haciendo cada vez menos posibles de satisfacer.

Por esta razón, Ostrom recomienda que no se confíe en una sola herramienta teórica para explicar el comportamiento humano (en este caso, solamente en la Teoría de la Elección Racional), a menos que deseen limitar sus análisis a situaciones que puedan ser modeladas como situaciones simples, competitivas y con información completa (p. 155). Al existir muchas situaciones relevantes, como en las que las reglas son objeto de elección (que es el caso de las situaciones objeto del presente estudio), que resultan muy complejas para ser analizadas como juegos formales, la utilidad de esta herramienta queda muy cuestionada (si bien, no del todo descartada). En general, se plantea que el concepto de comportamiento humano que emplea la Teoría de Juegos es limitado, y que es únicamente el extremo de un continuo de modelos apropiados para el análisis institucional (p. 44).

En todo caso, es importante denotar que los supuestos utilizados por la Teoría de Elección Racional (individuos poseen información completa sobre la estructura de la situación; individuos asignan una valoración interna completa y coherente a los resultados; y, los individuos tras analizar la situación, maximizan su beneficio neto en función de lo que esperan que hagan los demás) si son retomados como punto de partida para la construcción de componentes del comportamiento humano que el analista institucional debe tener en cuenta, y desde los cuales, se plantean los supuestos acerca de los participantes de una situación (p. 155). Sin embargo, esta construcción se realiza aumentado el grado de abstracción de los conceptos previos de la Teoría de la Elección Racional, haciendo de estos últimos, únicamente puntos particulares dentro de lo que viene a ser una tipología teórica. De este modo, lo relevante para explicar el comportamiento humano sería:

- La forma como los participantes adquieren, procesan, representan, retienen y usan la información.
- La valoración que los participantes asignan a las acciones y a los resultados.
- Los procesos que emplean los participantes para seleccionar acciones específicas o cadenas estratégicas de acciones a la luz de sus recursos.

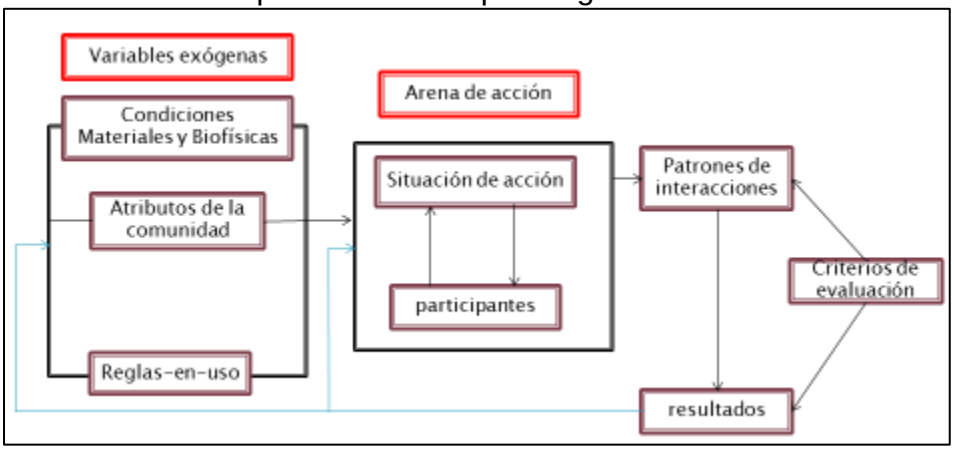
En este orden de ideas, considero que Ostrom tiene muchas ventajas por encima de las demás teorías para el análisis de políticas públicas, ya que además de contener las ventajas del neoinstitucionalismo económico de los 70’ y 80’ (como su claridad y coherencia conceptual, y el hecho de que consista en una teoría formal y deductiva, fundamentales para abordar teóricamente los objetivos específicos señalados), soluciona de una manera adecuada sus limitaciones referentes al uso desproporcionado de la Teoría de Juegos y de la Teoría de Elección Racional, así como su funcionalismo a las élites tecnocráticas al argumentar a favor de la democratización de las instituciones.

1.1.2.1. Esquema general IAD y múltiples niveles de análisis

El IAD sirve para hacer inferencias alrededor de los resultados que se obtienen en las políticas públicas, así como para ponerlas a prueba, lo cual es necesario para alcanzar el objetivo planteado. En este punto, Ostrom aclara: "En función de la estructura analítica de una situación y las hipótesis particulares que se empleen sobre el actor, el analista efectúa inferencias fuertes o débiles acerca de los resultados" (2015, p. 110).

Al analizar instituciones, se debe entender que sobre la base de los individuos se erigen estructuras compuestas por múltiples individuos, tales como familias, industrias o naciones, las cuales son, a su vez, partes de estructuras más amplias. Ostrom, siguiendo a Koestler, denomina a estos subconjuntos de unidades parte-todo anidados en sistemas adaptativos como "holones"[7]. La idea del IAD es diseccionar sistemas complejos en holones compuestos (Ostrom, 2015, p. 49). Lo primero y más fundamental, entonces, es identificar la unidad conceptual u holón denominado "arena de acción"[8], la cual puede ser utilizada para analizar, predecir y explicar el comportamiento de los actores en un arreglo institucional particular (Ostrom, 2007, p. 28), y, por ende, los resultados que surjan de este. La estructura general del modelo IAD de Ostrom describe estas relaciones:

Esquema N° 1: Esquema general IAD



Adaptado de: Ostrom, 2007, p. 27.

Este esquema identifica los tipos más amplios de variables estructurales que existen en los distintos tipos de arreglos institucionales. La conexión causal general del

[7] Es importante resaltar la equivalencia conceptual que tiene el concepto "holón" con el concepto "subsistema". Esto ayuda a entender los conceptos de "Patrón de interacción" y "Resultados" del esquema general del IAD, esencialmente como outputs de un sistema.

[8] Se refieren al espacio social donde los individuos interactúan, intercambian bienes y servicios, resuelven problemas, se dominan unos a otros o pelean entre sí -entre otras cosas que también hacen los actores- (Ostrom, 2007, p. 28).

esquema puede comprenderse del siguiente modo: la configuración particular de unas variables exógenas al incidir en la estructura de una arena de acción genera unas interacciones[9] de los cuales surgen unos particulares resultados (Ostrom, 2015, p. 51), estos últimos inciden de nuevo en el cambio, tanto de la arena de acción[10] como de las variables exógenas.

En este marco, la efectividad de la política pública de evaluación de políticas en Colombia corresponderá a los resultados que surgen desde los patrones de interacción, mientras que tanto la coordinación del proceso de evaluación con el de política evaluada, como la satisfacción de los estándares profesionales de calidad de evaluación corresponden a los patrones de interacción que se desprenden de la estructura de la arena de acción.

Esta estructura se puede dar en múltiples niveles, y por la misma razón, unas estructuras pueden estar anidadas en otras. Esto permite comprender la complejidad de las estructuras sociales, así como patrones de comportamiento y de resultados. Sin embargo, Ostrom aclara que más que una predicción precisa puntual, unos resultados coherentes con patrones será la mejor verificación que se pueda alcanzar en escenarios de alta complejidad (2015, p. 48), lo que en parte explica la selección del método comparado en esta investigación.

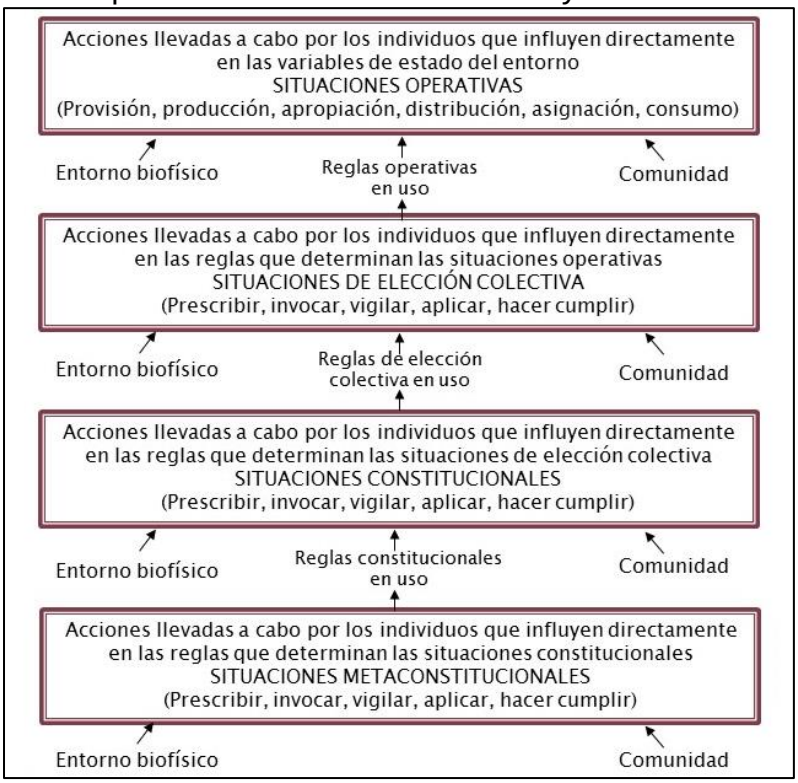
Plantear al IAD como un marco multinivel significa que las arenas de acción a identificar, así como sus problemas, pueden estar ubicados a nivel operativo, en donde los actores interactúan a la luz de los incentivos que enfrentan para generar resultados directamente en el mundo; a nivel de elección colectiva, donde los tomadores de decisión deben tomar decisiones de política pública -crear reglas para el nivel inferior- dentro de los límites de un conjunto de reglas de elección colectiva; o pueden estar ubicados a nivel constitucional, en donde las decisiones se dirigen a determinar quién es elegible de participar en la elaboración de políticas y acerca de las reglas que serán usadas en la elaboración de políticas (2007, p. 27) –crear reglas para elaborar otras reglas-[11]. Esta diferenciación, así como la interacción entre los distintos niveles, se esquematiza de la siguiente manera, en donde los tipos de resultados –por nivel- aparecen entre paréntesis en la parte inferior de cada nivel:

[9] Estas interacciones o “patrones de interacción” corresponden a los comportamientos de los individuos a los cuales las instituciones están dirigidas a modificar (Ostrom, 2015).

[10] Incide en estas variables ya sea en la misma arena de acción (del mismo nivel) o en las variables de otras arenas de acción (del mismo nivel o de niveles distintos como se verá más adelante).

[11] Considero relevante en este punto observar la similitud que tiene el considerar el cambio institucional dentro de los límites de las instituciones preexistentes, lo cual había sido considerado con anterioridad dentro del neoinstitucionalismo histórico.

Esquema N° 2: Niveles de análisis y resultados



Adaptado de: Ostrom, 2015, p. 103.

Es importante caracterizar todas las arenas de acción de los distintos niveles de la política pública de evaluación de políticas, sin embargo, en términos de efectividad en cuanto a impacto, siguiendo a Ostrom las más importantes vienen a ser las arenas correspondientes a situaciones operativas, puesto que son estas las que producen resultados que afectan directamente al mundo (variables de del mundo biofísico o de la comunidad) o a las demás arenas de acción (sean del mismo nivel o de otros niveles).

En nuestro caso de estudio, el muestreo gira entorno a los procesos particulares de evaluación desarrollados por el DNP, estos corresponden a arenas de acción de situaciones operativas, cuyo resultado es proveer la información que poseen los actores de otras arenas de acción. A continuación, se desglosan los componentes de las arenas de acción, y las variables que las explican.

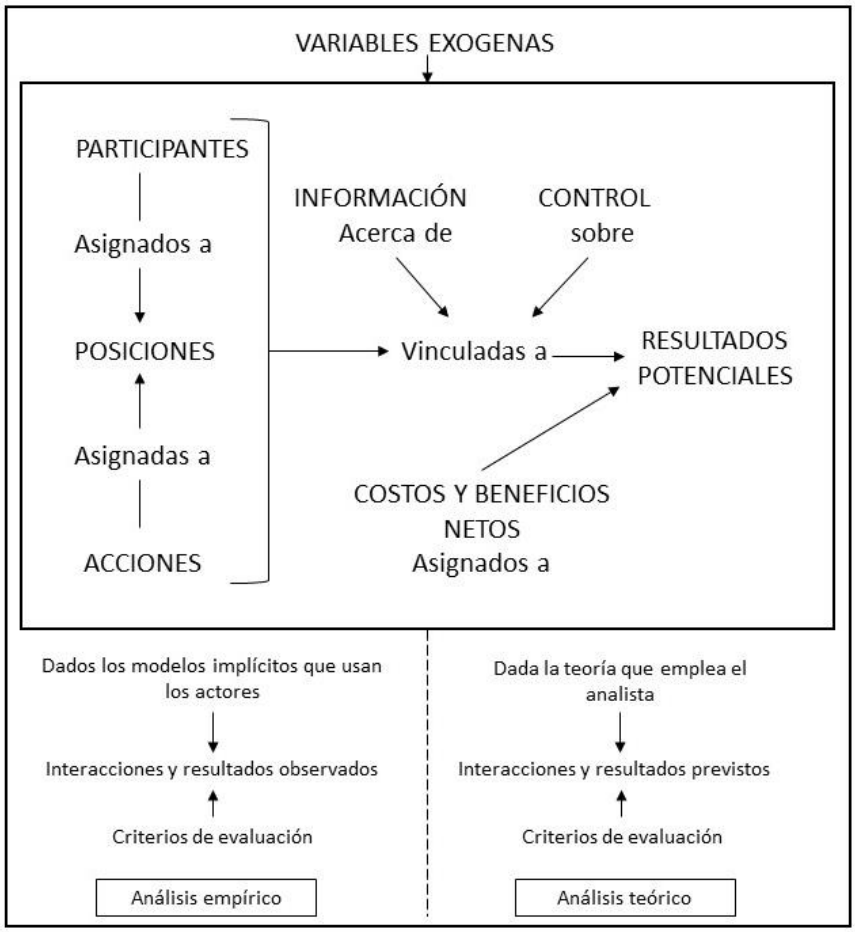
1.1.2.2. Situaciones de acción

Las arenas de acción incluyen como componentes una situación de acción y los actores (participantes) en esa situación de acción. Una situación de acción “Se refiere al concepto analítico que permite al analista aislar la estructura inmediata que afecta a un proceso de interés con el propósito de explicar las regularidades en

las acciones humanas y sus resultados, y el potencial de reformarlas (2007, p. 29). Esta, se caracteriza por tener siete agrupaciones de variables: el conjunto de participantes, sus posiciones específicas, resultados potenciales, las acciones posibles y sus vínculos con los resultados, el control ejercido por los participantes, información disponible a los participantes y los costos y beneficios asignados a los resultados (2007, p. 28).

La estructura interna de una situación de acción puede representarse con el siguiente esquema:

Esquema N°3: Estructura interna situación de acción



Adaptado de: Ostrom (2015, p. 74)

Esta estructura de la situación de acción tiene una particular dinámica de transformación, en la cual los individuos inciden de la siguiente manera:

> “Cuando se efectúa el análisis de una situación concreta, estas partes activas son los “datos” de la situación. Para poder analizar el comportamiento humano probable y los resultados dentro de una estructura particular se asume que la estructura de la situación es fija a corto plazo. Dentro de una situación particular, los individuos pueden intentar escoger solo a la luz de

> sus opiniones sobre la oportunidades y limitación de la situación (…) los individuos pueden, con el tiempo, influir en la estructura de las situaciones de acción en las que participan habitualmente modificando la configuración de reglas que influyen en la estructura de tales situaciones. Para hacerlo, se trasladan a niveles de análisis más profundos (situaciones de acción de elección colectiva o de elección constitucional), donde los resultados que se producen son cambios en las reglas de otras situaciones de acción" (2015, p. 74-75).

En cuanto a los componentes de la situación de acción, el conjunto de participantes determina quiénes y cuántos son (2007, p. 29). En esta parte de la situación se debe determinar ciertos atributos de los participantes –adquiridos o atribuidos-, que resultan relevantes para el análisis de la situación, entre estos atributos están: el número de participantes; su estatus como individuos, como equipo o como actor colectivo, y; otros atributos individuales como edad, educación, género y experiencia (2015, p. 80). Estas características influirán en las acciones y resultados, puesto que estos dependen de las habilidades y conocimiento que tengan los participantes (2015, p. 82).

Las posiciones llenadas por los participantes determinan qué roles existen y quiénes los desempeñan (2007, p. 29). Son espacios anónimos que constituyen el vínculo que conecta a los participantes con las acciones, puesto que en algunas situaciones cualquier participante puede llevar a cabo cualquier acción permitida en esa situación, pero en otras, la posibilidad de realizar particulares acciones queda asignada a posiciones concretas (2015, p. 83).

Los resultados potenciales son variables –sean cambios en el mundo biofísico, en la comunidad o institucionales- que están vinculados a secuencias de acción particulares (2007, p. 29). Para determinarlos es útil combinar los resultados biofísicos –transformaciones físicas subyacentes- de las acciones, las recompensas externas a los participantes por estos resultados biofísicos, y la valoración interna que hacen los participantes de estas al interior de una medida para analizar la toma de decisiones de estos (2015, p. 85). Al respecto, y con el propósito de avanzar en las distintas etapas del análisis es necesario tener en cuenta:

> "Cuando el objetivo del análisis de una situación se centra en comprender el resultado de una estructura particular y el analista está seguro del orden de los valores que los participantes asignan a los resultados, este proceso abreviado de representar los resultados en el espacio de utilidades constituye un medio eficaz para analizar la situación (…) Cuando el analista desea comprender cómo las reglas o características del mundo biofísico modifican una situación de acción, se logra mayor precisión analizando por separado los resultados materiales y posteriormente el valor que los participantes asignan a dichos resultados" (2015, p. 84).

Las acciones se refieren al conjunto de acciones admisibles y su vínculo con sus resultados potenciales (2007, p. 29). Estas acciones pueden ser concebidas como la selección de un estado o un valor en una variable de control, la cual, espera el participante que influya en la variable resultado (2015, p. 87). La acción específica que el participante selecciona dentro del conjunto de acciones permitidas a su posición se denomina “elección”, mientras que la especificación completa de todos los movimientos que se pueden efectuar en todas las contingencias de acciones posibles se denomina “estrategia” (2015, p. 88) –de los participantes-.

Las acciones posibles están vinculadas a particulares resultados. Ostrom plantea: “Un determinado valor en una variable de control se considera “vinculado” a una variable de estado –resultado- cuando es posible emplear dicho valor para hacer que la variable de estado: 1) vea la luz; 2) desaparezca, o 3) cambie de grado” (2015, p. 88). Esta es la definición puntual de causalidad que utiliza el IAD.

El siguiente componente de la estructura de la situación es el control que cada participante tiene sobre el vínculo existente entre acción-resultado. Aquí un participante puede tener control total sobre una variable resultado con su acción (probabilidad condicional es igual a 1), puede tener control parcial sobre una variable resultado con su acción en donde puede variar de tener una muy pequeña probabilidad de influir a tener una muy alta (probabilidad condicional mayor que cero y menor que 1), por último, puede tener un control nulo en donde el participante es impotente con respecto al resultado (2015, p. 93).

En cuanto a la información disponible para los participantes acerca de la estructura de la situación de acción se debe preguntar: ¿qué tanta información tiene los participantes acerca del recurso a gobernar, acerca de otros participantes, sobre funciones de costo/beneficios, y acerca de sus acciones que se agrupan en particulares resultados? (2007, p. 30). Existen situaciones en donde los participantes tienen información completa, esto es, cuando se asume que cada participante puede conocer toda la estructura de la situación de acción, tal como se ha definido (2015, p. 94).

Hay otras situaciones donde se tiene información incompleta (las cuales son más probables la mayoría de las veces), esto depende del modelo acerca del individuo, en donde pueden existir supuestos acerca de las limitaciones de la capacidad cognitiva humana (como por ejemplo la racionalidad limitada de Simon), así como de reconocer que:

> “La utilidad de la suposición de información completa o incompleta tiene también que ver con la complejidad de la situación de acción que se está estudiando y la estructura del lenguaje que se emplea, los canales de comunicación y el tipo de comunicación que, según se supone, suele estar presente en la situación de acción” (2015, p. 95).

El último elemento de la estructura de la situación son los costos y beneficios asignados a acciones y resultados. Estos, funcionan como incentivos y disuasivos,

y para analizarlos es bueno preguntar ¿qué tan costosas son las distintas acciones para los participantes? ¿qué tipos de beneficios se generan a partir de distintos resultados grupales? (2007, p. 30). En este punto es necesario distinguir entre los resultados materiales que obtienen o buscan los participantes, las recompensas o sanciones externas (que pueden asignarse a acciones, resultados, o a los vínculos entre estas) y la valoración intrínseca que hacen los participantes a los dos anteriores factores (2015, pp. 96-97).

Todos estos elementos de la situación de acción, tal como se observó en el esquema, están relacionados unos con otros. Estas relaciones Ostrom las sintetiza de la siguiente manera: Los participantes y las acciones se asignan a posiciones; los resultados están ligados a las acciones; la información está disponible a través de vínculos entre acción y resultado; el control se ejerce sobre los vínculos entre acción y resultado, y; a los vínculos entre acción y resultado se les asignan costos y beneficios (2015, p. 256).

1.1.2.3. Participantes

Teniendo en cuenta que la fuerza de las inferencias depende de las hipótesis particulares que se empleen sobre el actor, tal como se planteó arriba, los participantes o actores (individuales o corporativos) se caracterizan por cuatro agrupaciones de variables: los recursos que el actor trae a la situación; la valoración que los actores asignan a estados del mundo y a las acciones; el modo en que los actores adquieren, procesan, retienen y usan las contingencias de conocimiento y la información; y los procesos que los actores usan para seleccionar entre particulares cursos de acción (2007, p. 28).

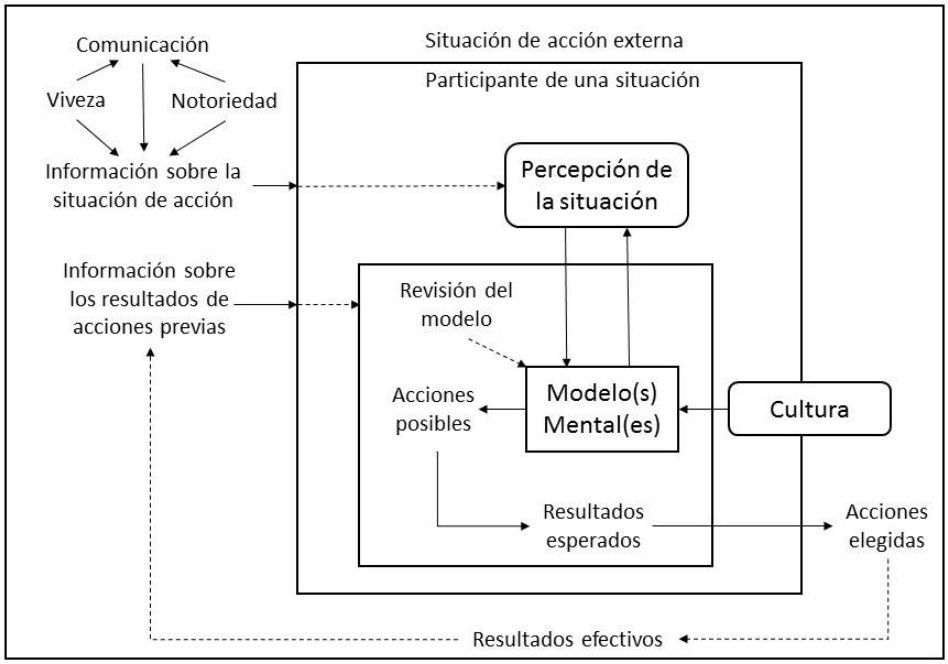
En cuanto al procesamiento de la información que hacen los participantes, Ostrom siguiendo a Holland (1986) plantea que los participantes crean “modelos mentales”, los cuales corresponden a representaciones de las diversas situaciones para ser capaces de tomar decisiones razonables en escenarios múltiples, estos modelos son determinados básicamente por la retroalimentación que hacen los actores del mundo, y por la cultura compartida o sistema de creencias en el que el participante se encuentra (2015, p. 157). Es de resaltar que Ostrom retoma el concepto de modelo mental de Holland, los cuales vienen a tener equivalencia con el sistema de creencias en Sabatier (2007) o con el denominado campo cognitivo conceptualizado por Muller (2002) en su análisis de referenciales de política, ambos autores pertenecientes al enfoque cognitivo para el análisis de políticas públicas.

Los participantes, para crear un particular modelo mental, comienzan por basarse en modelos mentales previos sobre la situación en la que se encuentran y de este modo calculan los beneficios esperados. En dado caso que estén satisfechos con los resultados no buscan más información, por el contrario, si no están satisfechos o detectan incongruencias, lo que aprende del resultado de sus acciones o la de los otros, le estimula para revisar sus modelos mentales (2015, p. 157). Puede decirse

que esta es la manera en la que Ostrom describe el papel del aprendizaje de los actores dentro de procesos de política pública, lo cual es necesario para explicar las razones por las que los actores de entidades sectoriales se inclinarían a utilizar la información producida por las evaluaciones para mejorar las políticas públicas evaluadas.

El siguiente esquema abre el holón “participante” dentro de la situación de acción, lo cual corresponde al modelo de individuo, en el cual, los individuos pueden describirse como limitadamente racionales, falibles pero adaptables (2015, p. 195), que persiguen objetivos múltiples para sí mismos y para otros, que adoptan normas de comportamiento que dependen del contexto, los cuales aprenden con el tiempo mejores estrategias para una situación concreta si se va generando información precisa sobre variables clave (2015, p. 173), y cuyo procesamiento de información, no responde inmediatamente a toda la información, sino a la que es comunicada de forma viva y notoria (2015, p. 161):

Esquema N°4: Participante en una situación de acción.



Fuente: Ostrom (2015, p. 162).

Es importante destacar que la búsqueda de información es costosa y las capacidades de procesar información de los humanos son limitadas, por lo que la mayor parte de las veces se toman decisiones con conocimiento incompleto de las alternativas posibles y de sus resultados probables (2015, p. 159). Gran parte de la utilidad de la revisión que se realiza a la teoría de la evaluación presentada más adelante es abrir el campo de posibilidades políticas y aumentar el rango de opciones factibles.

Frente al modo en el cual los participantes procesan la valoración que hacen de acciones y resultados, el supuesto básico que nos interesa es que el

comportamiento de estos está basado parcialmente en preferencias intrínsecas en cuanto al modo preferido de comportarse, los resultados que desean obtener ellos mismos, así como los resultados que esperan que otros obtengan (2015, p.167).

1.1.2.4. Variables causales de las arenas de acción

Tal como se observa en el esquema general IAD, los factores que afectan la arena de acción (y por lo tanto a los resultados que producirán) son tres agrupaciones de variables: las reglas usadas por los participantes para organizar sus relaciones, los atributos biofísicos del mundo[12] sobre los cuales son actuados en esas arenas, y la estructura de la más general comunidad en la cual la arena particular es situada[13].

Los atributos del mundo biofísico sobre el cual se actúa o se transforma determinan las acciones que son físicamente posibles, los resultados que pueden producirse, el modo en el que se vinculan las acciones a los resultados y la información que tienen los actores (2015, p. 62). Entre estos atributos, los más importantes son la exclusión, que tiene que ver con la dificultad de restringir el número de personas que se benefician de la provisión de un bien o servicio, y la sustracción, que se refiere a la medida en la que el uso por un individuo disminuye la disponibilidad del bien o del servicio para el consumo de otras personas (2015, p. 53). Al interior de la política de interés, el dinero con el que se contrata a los consultores (evaluadores externos), la información relevante, el tiempo, y el servicio de asesoría y facilitación, son recursos a los que se les deberá observar estas características en las distintas situaciones de acción relevantes.

En cuanto a los atributos de la comunidad se deben tomar en cuenta las normas de comportamiento generalmente aceptadas dentro de esta (2007, p. 43), el entendimiento común que tienen los potenciales participantes de la estructura de las distintas arenas de acción –aspecto sin el cual se dificulta cualquier coordinación- y el grado de homogeneidad en las preferencias de la comunidad[14].

---

[12] Ostrom especifica esta agrupación de variables del siguiente modo: "… some of the variables of an action situation are also affected by attributes of the physical and material world. What actions are physically possible, what outcomes can be produced, how actions are linked to outcomes, and what is contained in the actor's information sets are affected by the world being acted upon in a situation (…) The attributes of states of the world and their transformation are explicitly examined when the analyst self-consciously ask a series of questions about how the world being acted upon in a situation affects the outcome, action sets, action-outcome linkages, and information sets in a situation" (2007, p. 39).

[13] Los atributos de la comunidad que son importantes en la explicación de la arena de acción incluyen las normas de comportamiento generalmente aceptadas en la comunidad, el nivel de entendimiento común que los participantes potenciales comparten acerca de la estructura de particulares tipos de arenas de acción, el grado de homogeneidad en las preferencias de aquellos viviendo en una comunidad y la distribución de recursos entre los afectados (2007, p. 43).

[14] En este punto es importante tener en cuenta lo siguiente: "… when all appropriators from a common-pool resource share a common set of values and interact with one another in a multiplex set of arrangements, the probabilities of their developing adequate rules and norms to govern resources are much greater (…) If the

Por último, la distribución de los recursos entre los afectados deberá ser atendida en orden de sustentar algunas de las variables de la situación de acción como la factibilidad de los resultados buscados por los distintos actores.

Reglas son entendimientos compartidos entre aquellos implicados que se refieren a prescripciones impuestas acerca de qué acciones (o estado del mundo) son requeridos, prohibidos o permitidos. Todas son el resultado del esfuerzo implícito o explícito de alcanzar orden y predictibilidad entre humanos, a partir de crear clases de personas (posiciones) que son requeridas, permitidas o prohibidas para que tomen clases de acciones en relación con estados del mundo requeridos, permitidos o prohibidos (2007, p. 36). Complementando esto Ostrom plantea:

> "Los autores vinculados al análisis institucional emplean a menudo el término para denotar una regulación. La definición de reglas que empleamos en esta obra está próxima a lo que Black identifica en el sentido regulatorio. Se puede pensar en las reglas como un conjunto de instrucciones para crear una situación de acción en un entorno particular" (2015, p. 55).

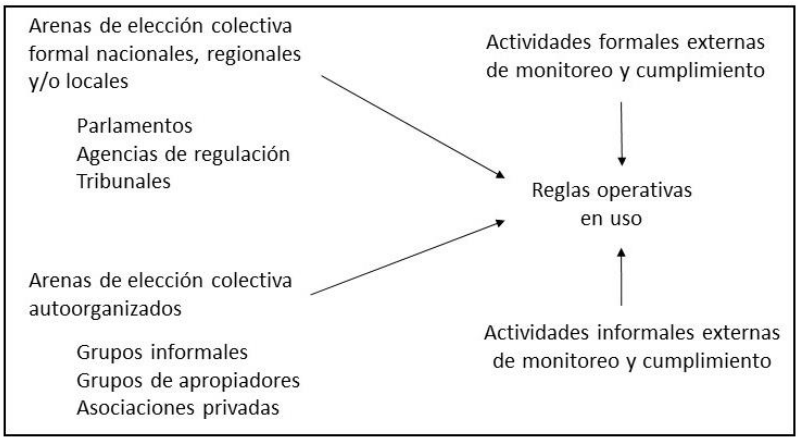
Cuando los individuos participan en la elaboración de distintas capas de reglas, algunas reglas quedarán establecidas de modo escrito (usando lápiz y papel en palabras de Ostrom, equivalente a que sean formales), pero muchas otras tomarán la forma de interacciones entre individuos para la solución de problemas y para descubrir un modo para hacer un mejor trabajo que en el pasado, por lo que los colegas en un equipo de trabajo crearan sus propias reglas (2007, p. 36), por lo que se establecerán algunas arenas de acción informales, que interactuarán con las formales para determinar las reglas en uso (p. 46).

Las relaciones entre arenas formales e informales se compendian con el siguiente esquema, que establece que además de que existen múltiples niveles de situaciones o arenas de acción, también pueden existir múltiples arenas en un mismo nivel, las cuales determinarán las reglas operativas en uso, que, al ser exógenas a la situación de acción y al provenir de arenas formales o informales, también pueden catalogarse como reglas operativas en uso formales o informales. En este punto podría considerarse la similitud que tienen los planteamientos de Ostrom, con los postulados del neoinstitucionalismo sociológico referentes a los elementos cognitivos interiorizados por los individuos y que determinan el comportamiento de los individuos.

---

appropriators from a resource come from many different communities and are distrustful of one another, the task of devising and sustaining effective rules is substantially more difficult" (2007, p. 43).

Esquema N°5: Relaciones entre arenas de elección colectiva formales e informales



Fuente: Ostrom (2015, p. 107)

Las reglas en uso[15] afectan todos los elementos de la situación de acción (2007, p. 39), por lo que, para identificarlas, es importante examinar cómo estas reglas afectan cada una de las variables en una particular situación de acción, este conjunto de reglas afectando a estas variables constituyen el mínimo, pero necesario conjunto de reglas para dar explicación a las acciones y resultados usados por los participantes para ordenar sus relaciones en una arena de acción (2007, p. 37).

Estas reglas no necesitan ser escritas, ni tampoco dimanar de procedimientos formales o legales, puesto que los individuos construyen reglas institucionales modificando la estructura de situaciones de acción reiteradas, con el fin de mejorar los resultados obtenidos (2015, p. 57). Ya sean escritas o no, lo importante es el significado que adquieren estas por parte de los participantes, puesto que de sto dependerá la estabilidad de los patrones de interacción:

> "La estabilidad de las acciones ordenadas por reglas depende del significado compartido que se asigna a las palabras que se emplean para formular un conjunto de reglas. Si no hay un significado compartido en el momento en que se formula una regla, existirá confusión en relación con las acciones permitidas, las obligatorias y las prohibidas" (2015, p. 59).

[15] Reglas-en-uso se refieren a las que son socializadas cada vez que alguien nuevo (como un empleado o un niño) entra en un existente sistema ordenado de reglas de comportamiento (2007, p. 23). El término "Working rules" se refiere al conjunto de reglas a las que harán referencia los participantes si se les pregunta que expliquen o que justifiquen sus acciones al resto de participantes" (p.36). También son: "las reglas a las que los participantes se remiten cuando las otras personas que participen en el juego les piden que expliquen o justifiquen sus acciones" (2015, p. 58).

De hecho, Ostrom aclara el papel fundamental que juega la comunicación en las distintas situaciones de acción, puesto que con esta se puede alcanzar una comprensión común de los problemas que se encaran de manera colectiva, se generan ideas respecto de los modos de lidiar con los problemas, y ayuda a los participantes a aprender normas y establecer la confianza necesaria para adoptar planes de acción basados en comportamientos fiables (2015, p. 110).

En el trabajo de laboratorio con instituciones (realizado por Ostrom y por muchos otros usando teoría de juegos), en muchos estudios han surgido conclusiones que establecen que incluso en entornos de extrema heterogeneidad entre los actores, la comunicación sigue siendo un mecanismo importante para establecer la coordinación entre actores[16] (2015, p. 137).

1.1.2.5. Gramática de las instituciones

Ostrom construye una herramienta de análisis gramatical de las instituciones, que sirve para resumir y analizar el contenido de las afirmaciones o enunciados institucionales[17], distinguir sistemáticamente entre los que se comprenden mejor como atributos de la comunidad (estrategias y normas), y los que se comprenden mejor como reglas que estructuran la situación de acción, y estudiar su formación y evolución (2015, pp. 196-197), frente a esto, se aclara puntualmente:

> "…afirmamos que las reglas están alejadas un paso gramatical de las normas y dos pasos de las estrategias. Reconocer explícitamente las diferencias entre reglas, normas y estrategias aclara varias cuestiones analíticas sobre las semejanzas entre estos conceptos y sobre la existencia y origen de las reglas institucionales. Mediante esta gramática, es posible plantearnos claramente cuestiones clave sobre cuándo y por qué las estrategias y las normas evolucionan hasta convertirse en reglas" (2015, p. 197).

La estructura gramatical o sintaxis de las instituciones incluye cinco componentes: 1) *Atributos*; 2) *Deóntico*; 3) *objetIvo*; 4) *Condiciones*; 5) *O de lo contrario*. Estos elementos constituyen los tres tipos de enunciados o afirmaciones institucionales, y los diferencia en que las reglas contienen los cinco componentes (ADICO), las

---

[16] En cuanto a la necesaria comunicación para coordinar actores también es relevante considerar: "Denzau y North (2000) subrayan la importancia que tiene la comunicación como instrumento para que los individuos desarrollen un modelo mental compartido. Si los mismos individuos interactúan repetidamente, comunicándose abierta y libremente, y si ya comparten amplias visiones culturales del mundo, es más fácil que sus modelos mentales converjan" (Ostrom, 2015, p. 159).

[17] Los enunciados institucionales se comprenden como: "un término de amplio alcance que comprende tres tipos de afirmaciones: reglas, normas y estrategias compartidas. Estos enunciados describen oportunidades y limitaciones que crean expectativas en el comportamiento de los demás actores (…) abarca un amplio conjunto de limitaciones y oportunidades lingüísticas compartidas que prescriben, permiten o recomiendan acciones o resultados a los participantes en una situación de acción. Subrayamos la naturaleza compartida que tienen estos conceptos. Muchos enunciados escritos tienen forma de regla (o de norma o estrategia), pero son desconocidos para los participantes y no influyen en el comportamiento. Estas afirmaciones se consideran reglas formales, pero no reglas en uso" (2015, p. 196).

normas contienen cuatro componentes (ADIC), y las estrategias, por su parte, solo contienen tres componentes (AIC). Lo cual proporciona un formato coherente para escribir afirmaciones institucionales, para que sean reescritas o resumidas así se expresen en lenguaje natural, y para analizar su evolución (2015, p. 2015).

En la presente investigación será pertinente observar si las reglas institucionales que conforman la teoría de evaluación (por ejemplo, los estándares de calidad), y que están relacionadas con las variables independientes, tienen el mismo estatus dentro del diseño institucional de la política pública analizada. En dado caso que el estatus fuera demasiado heterogéneo al comparar la política pública con la construcción institucional que existe en la teoría, no podría considerarse que la política sea un reflejo de la teoría de evaluación que se quiere poner a prueba.

Los “atributos” hacen referencia a las características particulares que deben tener los individuos o participantes a los cuales se les aplica la institución, por lo cual determina el subconjunto de participantes que se ven afectados por esta. Cuando no se especifican atributos particulares, el valor por defecto que va a tomar en el enunciado institucional es que el atributo incluye a todos los miembros en el grupo de participantes (2015, p. 200)

Por su parte, el componente “deóntico” se basa en las operaciones modales deónticas las cuales constan de definir lo que es permitido, lo que es obligatorio y lo que es prohibido. Estos se pueden encontrar en los enunciados institucionales en frases operativas como puede, debe/debería y no puede/no debería, las cuales son asignadas a acciones y a resultados (2015, p. 201). Es importante aclarar que estos operadores deónticos se vinculan a cada uno de los componentes de una situación de acción, lo cual hace surgir una particular clasificación de las reglas que será expuesta un poco más adelante.

El “objetivo” denota la descripción específica de una parte activa de la situación de acción (puede ser: participantes, posiciones, acciones, resultados, control, información o costos y beneficios) a la cual se refiere un enunciado institucional. Algunas veces, el objetivo se dirige a describir acciones que deben realizarse, otras, especifica ciertas situaciones del entorno o resultados en lugar de una acción (2015, p. 209), en ambos casos, para que el enunciado sea efectivo, tanto las acciones como los resultados deben ser posibles o evitables (tener ambas opciones desde la factibilidad que existe).

Las “condiciones” establecen el conjunto de variables que especifican cuándo y cómo se aplica el enunciado institucional. Si el enunciado no especifica variables particulares acerca de su aplicación, el valor por defecto de la “condición” va a ser en cualquier momento y en cualquier lugar abarcado por esa regla, norma o estrategia (2015, p. 210).

Por último, el “O de lo contrario” establece las consecuencias que un enunciado institucional asigna a la detección de un incumplimiento de algún otro de los componentes del enunciado. Es frecuente que con este componente se especifiquen un rango de posibles castigos cuando no se cumple la regla. Solo las reglas incluyen este componente, lo cual es fundamental para distinguirlas de los otros enunciados institucionales (2015, p. 211). Este componente debe ir acompañado de posiciones con suficiente autoridad como para vigilar y sancionar a quienes incumplan estas reglas, puesto que, si esto no existe, los enunciados institucionales que contengan un “de lo contrario” no llegan a tener una conceptualización como regla, tal y como la establece Ostrom (2015, p. 212), y por lo tanto no inciden como tal en el comportamiento humano en una situación determinada.

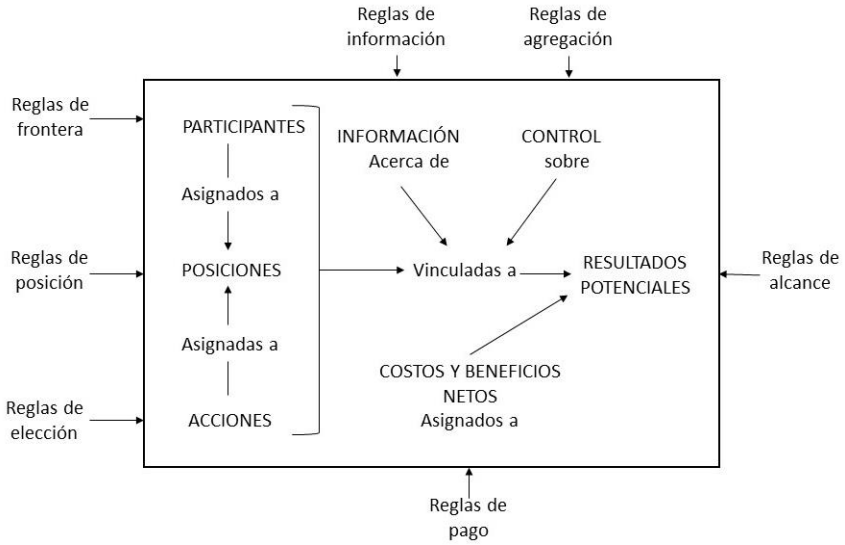
1.1.2.6. Clasificación de las reglas

Ostrom plantea una clasificación de las reglas alrededor del objetivo que tienen en su sintaxis, esto le permite “… identificar un importante subconjunto de todas las reglas que permita analizar dónde se han usado, qué otros factores influyen en los resultados deseados y que nivel de rendimiento han logrado” (2015, p. 249). Esto con el fin de establecer una clasificación que permita estudiar cuestiones concretas y examinar las reglas concretas que permiten abordar cada cuestión (2015, p. 250)

Los tipos de reglas son: 1) reglas de entrada y de salida[18]; 2) reglas de posición[19]; 3) reglas de alcance[20]; 4) reglas de autoridad[21]; reglas de agregación[22]; reglas de información[23]; y reglas de pago[24]. Estas reglas, como variables exógenas que influyen en todos los elementos de la situación de acción, se pueden esquematizar de la siguiente manera:

[18] Afectan el número de participantes, sus atributos, sus recursos, si pueden entran libremente, y sus condiciones para irse (p. 38).
[19] Establecen la posición en la situación -el rol o el tipo de participante que será- (p. 38).
[20] Delimitan los resultados potenciales que pueden ser producidos, y moviéndose hacia atrás, las acciones vinculadas a esos resultados específicos. Determina por ejemplo el dominio geográfico o funcional que es permitido a prohibido (p. 38).
[21] Asignan un conjunto de acciones que los participantes en las posiciones particulares pueden, deben o no deben hacer. Estas al estar combinadas con leyes científicas acerca de los estados del mundo, determinan el árbol de decisiones que pueden tomar, esto es: los vínculos acción-resultados. Determina acciones individuales como por ejemplo las tecnologías de cosecha permitidas, autorizadas o prohibidas (p. 39).
[22] Afectan el nivel de control que los participantes en una posición ejercen en la selección de una acción. Por ejemplo: ¿ciertas acciones requieren el permiso o el acuerdo de otros actores? (p. 39).
[23] Afectan la información del conocimiento contingente de los actores, como por ejemplo reglas que determinan la información que debe ser pública o la que debe ser secreta (p. 39).
[24] Afectan los costos y los beneficios que son asignados a particulares combinaciones de acciones y resultados, de modo establecen los incentivos y disuasivos para los actores (p. 39).

Esquema N°6: Reglas como variables exógenas actuando sobre la situación de acción



Adaptado de Ostrom (2015, p. 256).

1.1.3. La evaluación de políticas públicas como disciplina pragmática

Actualmente, se entiende por evaluación de políticas públicas a la determinación del valor de un proyecto, programa o política. Más puntualmente, se refiere al proceso de determinar el valor o significancia, tan sistemático y objetivo como sea posible, de una planeada, en curso o completada intervención. Su propósito último es proveer información a los decisores para que puedan tomar mejores decisiones en una intervención (en lo cual radica su utilidad), ayudándolos a entender lo que es probable que suceda, está ocurriendo o ha ocurrido a causa de una intervención e identificar maneras de obtener más de los beneficios deseados. Es una valoración periódica y busca responder a la pregunta ¿Por qué? (Morra y Rist, 2009, pp. 8-16).

Existen múltiples modelos y diversas clasificaciones de la evaluación, cuyo criterio para distinguirlos algunas veces corresponden al énfasis del valor que se utiliza (metas, impactos, productividad, insumos, procesos, etc.), al enfoque investigativo o metodológico de la evaluación, a la relación que tiene el evaluador con respecto a la intervención evaluada, o al momento en que se realiza la evaluación. Es por esto que no existe un único criterio para clasificar los modelos de evaluación (Solarte, 2004, p. 100).

El Banco Mundial en su Informe Sobre el Desarrollo Mundial de 1997, se plantea el desafío del rediseño y reconstrucción del Estado (a pesar de haber estimulado su desmantelamiento poco tiempo atrás), el cual debe asumirse desde criterios de la racionalidad funcional, racionalidad operativa, en la eficiencia y en transparencia del desempeño (Ramírez, 2009a, p. 123-124). Esto es sustentado en el paradigma de

reinvención del gobierno de Osborne y Gaebler (1994, original de 1992), en donde se plantea que un gobierno puede funcionar como una empresa privada, si se sustituye una burocracia basada en el cumplimiento de reglas, por una institucionalidad basada en el control de resultados, el cual, en lo que respecta a las políticas públicas, tomará la forma de evaluación y seguimiento.

Teniendo en cuenta que la variable dependiente de la presente investigación es la efectividad de la política pública de evaluación de políticas en Colombia, se debe considerar que teóricamente existe un amplio consenso acerca del carácter pragmático de la evaluación de políticas, en la medida en que se enfoca en el mejoramiento de los procesos evaluados (políticas, programas y proyectos), esta es su utilidad y lo que define su finalidad. Para el logro de este fin, es obligatorio retroalimentar la información para una posterior toma de decisiones. Frente a esto Weiss plantea: *"*la evaluación en su calidad de investigación aplicada está consagrada al principio de la utilidad. Si no tiene ninguna influencia en las decisiones resultará una tarea en vano." (1975, p. 23).

En consonancia, Kusek y Rist manifiestan que una evaluación "debe proporcionar información creíble y útil, permitiendo la incorporación de lecciones aprendidas en el proceso decisorio" (2005, p. 12). Álvarez en el mismo sentido establece: "En la evaluación se determinan los efectos observables de las acciones emprendidas y se mide la distancia entre los logros efectivamente alcanzados y el objetivo previsto, lo cual retroalimenta la fase de optimización de la subsiguiente decisión" (1992, p. 17).

Los resultados de la evaluación pueden usarse de varios modos: dar a los gerentes y hacedores de políticas información sobre lo que funciona bien y mal en relación con los objetivos; ayudar a la localización de recursos; repensar las causas de problemas; identificar problemas emergentes; aportar la toma de decisiones entre alternativas contendientes; apoyar la reforma e innovación en el sector público; construir consenso en las causas de un problema y cómo responder a estos (Morra y Rist, 2009, pp. 14-15). Tal como se puede observar, el uso de la evaluación constituye la finalidad de esta, e independientemente del modo en que se manifieste el uso de los resultados de las evaluaciones, su esencia radica en la racionalización de la toma de decisiones a partir de información o evidencia.

Como se verá en el próximo capítulo, el objetivo de los diversos procesos evaluativos que se realizan en el marco de la política pública de evaluación de políticas es exactamente igual a esta finalidad ampliamente consensuada en el terreno académico. Teniendo en cuenta que la efectividad es la medida del grado en el que una actividad alcanza sus objetivos[25] (Morra y Rist, 2009, p. 29), podemos decir que, para fines de operacionalización de la variable dependiente de la presente

[25] Al medir la efectividad es útil considerar las siguientes preguntas: ¿en qué grado fueron alcanzados los objetivos/ o son probables de ser obtenidos?; ¿Cuáles son los principales factores que influencias el logro o fracaso de los objetivos? (Morra y Rist, 2009, p. 542).

investigación, el uso de la evaluación es un factor de medición esencial. Ahora la pregunta es ¿qué factores determinan que una evaluación se utilice o no?

1.1.3.1. Debate epistemológico acerca de evaluación de políticas públicas

Para responder a la anterior pregunta primero se debe contextualizar un breve recorrido histórico frente a lo que ha sido la evaluación como investigación aplicada, y detallar las discusiones de las cuales surge la actual teorización normativa. Frente a lo que se podría denominar la discusión espistemológica, Guba y Lincoln consideran la existencia de cuatro generaciones de evaluación, que se distinguen dependiendo del significado que se le ha atribuido a la evaluación[26] durante el siglo previo a su texto (1989, p. 22). Este es un texto relativamente antiguo, pero que caracteriza muy bien la discusión actual entre dos de los principales paradigmas de las ciencias sociales frente a la evaluación.

La primera generación se enmarca en el enfoque positivista, la segunda y la tercera dentro del postpositivismo[27], el cual se caracteriza por privilegiar métodos cuantitativos, la investigación experimental, el análisis estadístico y la utilización de instrumentos confiables (Solarte, 2004, p.79). Desde esta postura epistemológica, considera que las conclusiones sobre determinada intervención deben establecerse en función de evidencias objetivas, las cuales pueden ser obtenidas a partir de métodos experimentales (Solarte, 2004, p.83). Por su parte, la cuarta generación es la propuesta de Guba y Lincoln, la cual es fundamentalmente constructivista.

La primera generación de la evaluación es el resultado de un proceso de desarrollo de varias influencias. La primera es la necesidad de mediciones de varios atributos de los niños estudiantes de colegios (Guba y Lincoln, 1989, p.22), lo cual derivó en el desarrollo del test de coeficiente intelectual (IQ test), creado por el psicólogo Alfred Binet con el propósito de separar (clasificar) los niños “normales” de aquellos que presentaban deficiencias cognoscitivas:

> “Another application of testing that was to have major import occurred in France. The French national minister of education, harried by teachers demanding that he find some way to screen out mentally retarded youngsters,

[26] Significados atribuidos que reflejan el contexto histórico existente, los propósitos que la gente tiene al momento de hacer evaluaciones, y los supuestos filosóficos del evaluador (1989, p. 22).

[27] Solarte al describir esta corriente manifiesta: “El enfoque post positivista es una reacción al positivismo clásico y reconoce que existe una realidad objetiva, por fuera del observador, esta solo puede ser comprendida de manera imperfecta y probalística. De esta manera, la realidad no podrá ser entendida en su totalidad, aunque sea posible aproximarse a ella. Los postpositivistas plantean la necesidad de aproximarse a la realidad desde diversas perspectivas y defienden métodos experimentales modificados, aunque complementados con métodos cualitativos buscando, más que comprobar hipótesis, falsearlas. Los postpositivistas son realistas críticos.” (2004, p.78).

> who, it was said, were making it impossible to teach “normal” children (a kind of reverse mainstreaming, as it were), asked a psychologist, Alfred Binet, to devise a test for that purpose (…) Binet devised a new approach, based on the commonsense observation that mentally retarded youngsters would not be able to cope whit simple life situations such as counting money or identifying household objects as well as their normal counterparts.” (1989, pp. 23-24).

La segunda influencia es la necesidad de clasificar a partir de mediciones a los integrantes del ejército estadounidense, en el contexto de la Primera Guerra Mundial, la cual surge de evidenciar la gran utilidad de los exámenes de coeficiente intelectual desarrollados por Binet, y fue la que particularmente propició el desarrollo de estos procesos (p. 24).

Una tercera influencia es el gran desarrollo de las ciencias sociales, las cuales respondieron al llamado de John Stuart Mill por la aplicación de un “enfoque científico”, al interior del estudio de los fenómenos sociales y humanos, el cual era fundamentado en la confiabilidad del enfoque, y en el desarrollo de la física y la química en el siglo XIX (1989, p. 24-25). Esto generó la adscripción de la psicología en este enfoque epistemológico, que intentará emular las ciencias naturales en la medida de lo posible:

> “Psychology in particular became wedded to the scientific approach, attempting to emulate the physical sciences as closely as possible. Of course, this intent was well served by the apparently precise quantitative measurements that test were yielding.” (1989, p. 25).

La segunda generación de la evaluación se desarrolló para dar respuesta a la que en su momento se identificó como la principal falencia de la primera generación: tomar como objeto de la evaluación a los estudiantes de los colegios, o a las personas en general. Luego de la primera guerra mundial, se evidenció que los programas escolares necesitaban una revisión drástica, pero el enfoque de la evaluación, al proporcionar solo información sobre los estudiantes, no servía a este propósito (1989, p. 27).

El nacimiento de la evaluación de programas es obra del estadounidense Ralph Tyler, quien desde la Ohio State University trabajó para desarrollar exámenes que pudieran medir si realmente, los estudiantes aprendían lo que los profesores intentaban enseñarles, y así perfeccionar los programas educativos (hoy evaluación formativa), esto lo hacía recolectando información sobre el logro de los objetivos del programa, y con un análisis centrado en la descripción de fortalezas y debilidades (1989, p. 28). El papel del evaluador era entonces básicamente el de un descriptor (lo que no estaba en contraposición con el anterior enfoque centrado en la medición).

La tercera generación de la evaluación se origina también para solucionar problemas identificados en la anterior generación. En determinado momento se

esperó que la evaluación no fuera únicamente descriptiva, sino que también valorará según mérito, fortalezas y debilidades. Este enfoque valora el proceso, el rendimiento y los objetivos y metas del programa (1989, p. 30). Es por esto, que lo característico de esta generación, va a ser la inclusión del juicio normativo (en contraste con algún estándar o norma), sin abandonar la medición y la descripción:

> "The call to include judgment in the act of evaluation marked the emergence of third generation evaluation, a generation in which evaluation was characterized by efforts to reach judgments, and in which the evaluator assumed the role of judge, while retaining the earlier technical and descriptive functions as well (…) First, it required that the objectives themselves be taken as problematic; goals no less than performance were to be subject of evaluation" (1989, p. 30).

Guba y Lincoln plantean la existencia de tres dificultades principales al interior de estas tres generaciones: la tendencia hacia el gerencialismo, la falla al introducir un pluralismo de valores, y un compromiso excesivo con el paradigma científico de investigación -tradicional- (1989, p. 32). Las cuales, en esencia, son críticas hechas desde el constructivismo a la evaluación postpositivista actualmente dominante en el campo.

La tendencia hacia el gerencialismo trata de los riesgos latentes en la relación existente entre el evaluador y el contratante de la evaluación o gerente, en esta relación es este último quien tiene la última palabra al momento de ajustar parámetros y límites de la evaluación (1989, p. 32). El gerente o contratante de la evaluación (puede ser un actor colectivo o individual) define las preguntas que la evaluación pretende contestar, por lo tanto, induce las respuestas que se van a dar, a interpretar, y, además, la difusión de los resultados (1989, p. 32).

La falla -relacionada con la introducción de un pluralismo de valores- se relaciona con el cuestionamiento acerca de la idoneidad de la postura epistemológica objetivista y la neutralidad de los valores que guían la investigación científica (positivista) aplicados en la investigación evaluativa. Guba y Lincoln resaltan que los hallazgos son sujeto de múltiples interpretaciones, y los hechos que pueden ser encontrados, a su vez, son determinados por el sistema de valores que el evaluador tiene, aun si este lo desconoce (1989, p. 35). Al ser personas, con un particular sistema de valores, quienes desde este definen los límites y parámetros de la evaluación (controlando así las preguntas, la metodología y los resultados de la evaluación), no se puede afirmar que se están describiendo los fenómenos evaluados tal como son (1989, p. 35) o, dicho de otro modo, de manera objetiva, o, por lo menos, neutral.

Por último, el compromiso excesivo con el paradigma de investigación científica –tradicional- surgió de la pretensión de que los científicos sociales fueran sistemáticos y racionales en su pensamiento, tal como sucedía en las ciencias

naturales, también por el deseo de alcanzar una legitimidad de las disciplinas, siguiendo parámetros de rigurosidad metodológica tanto como fuera posible (1989, p. 35). Guba y Lincoln exponen resultados desafortunados de este excesivo compromiso.

Uno de estos es la separación que se hace del objeto de estudio con el contexto en el cual se desenvuelve. Esto responde al interés por aumentar rigor (o validez interna) a partir de estudiar el objeto de evaluación sin la interferencia de las variables de contexto en las variables de estudio. Frente a esto, los autores subrayan que las generalizaciones son imposibles en este sentido (1989, p. 36), puesto que los resultados de estas investigaciones únicamente reflejan lo que sucede bajo estas específicas condiciones experimentales.

Este compromiso excesivo con el paradigma de investigación científica -tradicional-, fundamentado en requerimientos de fortaleza y rigor de los datos que se esperan obtener del proceso de la evaluación, derivó en una dependencia de las mediciones cuantitativas. Conceptos como Información o datos fuertes han implicado básicamente datos cuantificables, puesto que estos pueden ser medidos con precisión con herramientas matemáticas y estadísticas (1989, p. 37).

Los evaluadores al utilizar este tipo de datos, por lo general, aseguran a sus clientes que la información de sus investigaciones corresponde a la realidad misma, por lo que no puede existir discusión o negación de lo que esta manifiesta. Esta situación, frecuentemente, desconoce la influencia de los valores, los clientes, o distintas situaciones en los resultados presentados. Este compromiso excesivo con el paradigma de investigación científica tradicional -entonces- refuerza y sostiene las tendencias hacia el gerencialismo (1989, p. 37) en los procesos de evaluación de política pública.

Frente a este punto, Solarte establece: “Las políticas son “configuradas” antes que “hechas”, lo cual desvirtúa el modelo de toma racional de decisiones en el sistema. La democracia “evoluciona para proteger de la tiranía” y no para impulsar modelos racionales; la toma de decisiones racionales está relacionada con el totalitarismo en la medida en que se aleja del ordenamiento de intereses de stakeholders el cual se logra mediante procesos políticos (2004, p. 95).

Estas fallas de las tres generaciones de la evaluación fundamentan en Guba y Lincoln el surgimiento de su propuesta de una Cuarta Generación de Evaluación (enmarcada en el paradigma constructivista), con la cual se establece el debate contemporáneo en contraposición a los postulados de la corriente postpositivista. Este debate Solarte lo describe de la siguiente manera:

> “Su discusión, que ha sido –y sigue siendo- una parte fundamental del debate evaluativo, radica en el origen del conocimiento y su naturaleza, lo cual equivale a la cuestión ontológica sobre la naturaleza de la realidad, epistemológica acerca de sus características y metodológica acerca de cómo construir conocimiento (…) La discusión en el terreno evaluativo se ha dado

> básicamente entre los postpositivistas y constructivistas, que corresponden a dos de los paradigmas básicos de la investigación en las ciencias sociales." (Solarte, 2004, p.77).

La postura postpositivista sostiene que existe una realidad gobernada por leyes universales, sin embargo, es imposible percibirla desde los mecanismos sensoriales e intelectuales imperfectos de los seres humanos (ontología: realismo crítico). La objetividad es el ideal regulatorio que se busca a partir de la neutralidad frente al objeto de estudio, junto con el diálogo crítico con la comunidad académica (epistemología: objetivismo modificado). De lo anterior se desprende el uso crítico de múltiples métodos, fuentes, datos y teorías que deben ser triangulados para verificar hallazgos (metodología: multiplismo crítico), incluyendo métodos cualitativos (Guba, 1990, pp.20-23).

El constructivismo, por su parte, sostiene que la realidad tiene la forma de múltiples construcciones mentales, basadas en la interacción social y la experiencia (ontología: relativista), niega la posibilidad de un dualismo entre el sujeto y el objeto, estableciendo que el conocimiento sobre el objeto existe por la interacción entre el observador y lo observado, y se encuentra precisamente dentro la mente de los sujetos (epistemología: subjetivismo), por lo cual, para adquirirlo, se deben interpretar las "realidades" de los sujetos, y posteriormente se deben poner a dialogar estas posturas, para así desarrollar una nueva construcción lo más informada y sofisticada posible (metodología: hermenéutica-dialéctica) (Guba,1990, pp. 25-27).

Es así como en la Cuarta Generación de Evaluación "la verdad" será cuestión de consenso entre los interesados de la evaluación, "los hechos" solo tienen significado al interior de un particular marco de valores, "las causas" y "los efectos" solo pueden ser imputados, "los fenómenos" solo pueden ser entendidos en su contexto, "la evaluación" produce información solo a partir de valores, por lo que valorar es una parte esencial de esta al atribuir cualquier significado, "los evaluadores" deben ser orquestadores del proceso de negociación, haciendo que culmine en un consenso a partir de información más sofisticada (1989, p. 44-45).

Sin embargo, esta posición ha recibido críticas por parte del postpositivismo, entre las cuales se encuentra la ambigüedad existente, la falta de control experimental de variables, y la dificultad de encontrar soluciones amplias y prácticas para problemas acuciantes. Guba y Lincoln se defienden de estas críticas al definirlas también como construcciones, en las cuales se han atrapado sus constructores por su rígida adherencia a supuestos, en los cuales se basa su credibilidad (Guba y Lincoln, 1989, p. 47). Sin embargo, la discusión parece estar muy lejos de llegar a una conclusión.

Esta investigación concuerda con la postura de Patton al establecer que uno no puede, de modo razonable, preguntar cuál marco teórico es "correcto", "mejor", o

definitivamente más útil que los demás, puesto que esto depende de lo que uno desee hacer y de cuáles supuestos uno comparta (2002, pp.134-135). Por esta razón, tanto la evaluación postpositivista como la constructivista tienen fortalezas, debilidades particulares, y deben ser juzgadas con diferentes criterios.

Al centrarse en los criterios o estándares adecuados para juzgar la calidad y credibilidad de una evaluación postpostivista, Patton resalta: objetividad (intento de reducir sesgo), validez de los datos, rigor sistemático en procedimientos de trabajo de campo, triangulación (consistencia de los hallazgos a través de métodos y fuentes de datos), confiabilidad de códigos y análisis de patrones, correspondencia de los hallazgos con la realidad, generalización (validez externa), fortaleza de las evidencias apoyando hipótesis causales y contribución a la teoría (2002, p. 544).

Por otra parte, los criterios o estándares para juzgar la calidad y credibilidad de una investigación (o evaluación) constructivista son: reconocimiento de subjetividad (discute y toma en cuenta sesgos), integridad (rigor), autenticidad y triangulación (capturar y respetar las distintas posturas), reflexividad y praxis (autocrítica de propio contexto y consideración de experiencias que afectan estudio), particularidad (integridad de la singularidad de los casos), entendimiento profundizado y mejorado, y contribución al dialogo (2002, pp. 544-546).

1.1.3.2. Teorizaciones recientes acerca de los factores que inciden en la efectividad/uso de la evaluación

Existe un interés creciente en la literatura acerca de evaluación de políticas públicas sobre el uso de la información y sus explicaciones, lo que significa un cambio en la orientación de esta literatura, en la cual prevalecía la discusión de las características que debían tener los sistemas de seguimiento y evaluación para producir información confiable, válida y relevante. Con respecto a las teorizaciones que se han realizado sobre el uso de la información, autores como González (2016), Johanson y Siverbo (2009), y Saliterer y Korac (2013) establecen que actualmente explicaciones que especifican los factores que inciden en el uso de la información pueden clasificarse como económicas-racionales, sociológicas y multiteóricas. Los siguientes autores se pueden ubicar en alguna de estas tres corrientes.

En cuanto a la primera corriente, el racionalismo, muy de la mano con el neoinstitucionalismo económico descrito anteriormente, se basa en las siguientes hipótesis: la idea de que los individuos pueden usar la información para tomar mejores decisiones, lo que en consecuencia generará un rendimiento mejorado (Saliterer y Korac, 2013, p. 2); la realización de evaluaciones crea las condiciones necesarias para la eficiencia a partir de reducir las asimetrías de información y los costos de transacción, ubicando de manera óptima los recursos, identificando servicios improductivos, estimulando la competencia entre los agentes y mejorando el flujo de la información (Johanson y Siverbo, 2009, p. 2); la medición del desempeño a través de indicadores viene acompañada invariablemente de

actividades de uso de la información para la toma de decisión (Taylor, 2011, p. 1); y que el uso de la información para la toma de decisiones puede ser facilitado principalmente a partir de avances estructurales (institucionales) y técnicos en las mediciones (p. 1).

En esta corriente surge la idea de que todo lo que los agentes públicos necesitan para hacer sus decisiones administrativas de manera competente, son algunos estándares o medidas de la calidad de los servicios y conocimiento sobre los costos de las diversas opciones (Ammons y Riverbank, 2008, p. 6). Estos estándares de calidad corresponden a arreglos institucionales, los cuales imponen restricciones a la acción de los actores con respecto a la planeación y la implementación de las evaluaciones, brindando opciones para realizar acciones que influencien la utilización de las evaluaciones (Balthasar, 2009, p. 252).

Siguiendo esta línea de investigación, Taylor (2011, p. 4) manifiesta que la calidad del sistema de medición del desempeño de una organización está, significativamente, relacionado con el uso de la información para la toma de decisiones. Esto lo hace, apoyándose en el enfoque estructuralista, en donde se enfatizan reglas formales, procedimientos y rutinas para la obtención de aprendizaje organizacional, y donde la importancia de los buenos procedimientos y rutinas es extendida a la calidad del sistema de medición (p. 3). Este autor manifiesta que, si bien muchos autores recientes se han enfocado en la importancia de otros factores a parte de los técnicos como explicativos del uso de la información para la toma de decisiones, también es cierto que estos autores no afirman que esos otros factores sean más importantes que los técnicos (p. 1), como pasa con Julnes y Holzer (2001), o con Sandison (2006).

Esta última autora, establece en el mismo nivel de importancia factores organizacionales, relacionales usuario-evaluadores, externos (contexto), y de calidad del sistema de medición del desempeño (pp. 101-102), la cual dependerá del modo en que la evaluación es diseñada, planeada, conducida y administrada por el cliente y el evaluador. Profundizando más en el asunto, Sandison aclara factores de calidad en materia de propósito de la evaluación, diseño, participación de actores y usuarios relevantes, planeación de la utilización, calidad de la evidencia, factibilidad y relevancia de las recomendaciones, mecanismo para el seguimiento de la utilización de la evaluación, y la credibilidad del evaluador (pp. 108-109). Otros autores de la misma corriente han subrayado la importancia de factores institucionales que influencian la distancia y la relación entre los evaluadores y los evaluados, y que los usuarios sean los responsables del costo de la evaluación (Balthasar, 2009).

Para Julnes y Holzer la perspectiva racionalista tiene como supuesto que la Al profundizar en los factores racionales que explican la utilización de estas mediciones, resaltan la importancia de comprometer los recursos necesarios; información sobre el modo de implementar las medidas del desempeño; orientación

hacia el cumplimiento de metas; requisitos externos a la organización sobre el uso de las mediciones (pp. 2-3); realizar una valoración de la preparación de la organización para desarrollar e implementar mediciones de su desempeño; identificar e involucrar a los distintos grupos de interés; y por último, enfatizar la necesidad de desarrollar una cultura de mejoramiento del desempeño (p. 11).

Otra perspectiva similar es la de Berman y Wang (2002) quienes se enfocan en la importancia que tiene una cuidadosa atención de las capacidades administrativas y organizacionales para una exitosa implementación de las medidas de desempeño (p. 2). Como dimensiones de esta capacidad, concluyen argumentando a favor de la importancia que las jurisdicciones sean capaces de relacionar resultados a las operaciones, recolectar datos a tiempo, tener un personal capaz de analizar datos sobre el rendimiento de la organización, una adecuada tecnología para el manejo de información, y apoyo del nivel directivo y de los oficiales elegidos (p. 10).

Siguiendo la misma corriente racionalista, Van de Walle y Van Dooren (2010) subrayan la importancia que tiene superar los marcos o umbrales organizativos y psicológicos para percibir, interpretar y usar la información, especialmente la que no les gusta ver (p. 19); la focalización en la organización y en las personas que deben usarla; capacidad y voluntad para actuar sobre la información (p. 20); calidad basada en verosimilitud y utilidad de la información, y la capacidad de la información para dilucidar y resolver problemas (p. 21).

Dentro de la discusión sobre el uso de la información de las evaluaciones, esta perspectiva ha recibido críticas que están en consonancia con las que se le hacen al neoinstitucionalismo económico en general. Algunas de estas son: una mayor cantidad de información no necesariamente reduce la incertidumbre, y hace poco para mejorar el entendimiento que se tiene de los problemas sociales; adicionalmente, para que la información tenga un impacto sobre la toma de decisiones necesita más que existir, por lo que el uso de la información es un proceso ambiguo, amorfo, incremental, incoherente y no ceñido a rutinas (Van de Walle y Van Dooren, 2010).

Respecto a la corriente sociológica, los autores explican el uso de la información a partir de elementos como el marco institucional, la cultura organizacional, el contexto de la organización, y sus procesos internos (Gonzáles, 2016, p. 229). Moynihan, Pandey y Wright (2012) en un texto enfocado hacia el desarrollo de un modelo conceptual que conecte el liderazgo transformacional con el uso de la información del desempeño, subrayan la importancia de una serie de variables intermedias que logran conectar dichos elementos, dentro de estos se encuentran: claridad acerca de las metas, cultura organizacional del desarrollo, disponibilidad de información acerca del desempeño (p. 15). En un texto previo, Moynihan y Pandey (2010) argumentaron a favor del papel que juega la motivación organizacional, el liderazgo, la disponibilidad de información, la cultura organizacional y la flexibilidad administrativa para la utilización de la información del desempeño. Otro autor que resalta el papel del liderazgo como factor explicativo del uso de la información es Askim (2007).

Siguiendo con las explicaciones sociológicas, Bourdeaux y Chikoto (2008) se adentran en los factores políticos y contextuales que indicen en el uso de la información para la implementación de reformas administrativas. En este marco de análisis, examinan particularmente la influencia de las organizaciones legislativas en el uso administrativo de las medidas de desempeño (p. 1). Este estudio concluye que las legislaturas ciudadanas se asocian mejor que las legislaturas profesionales a mejores prácticas administrativas. Adicionalmente, la heterogeneidad y la fragmentación política, y el diálogo dinámico entre los partidos políticos, y la supervisión legislativa a las organizaciones administrativas del sector público también sobresalieron como factores que fomentan la implementación de reformas a partir de mediciones del desempeño.

Finalmente, en cuanto a los modelos explicativos multi-teóricos, Johansson y Siverbo (2009), en un estudio orientado al examen de las explicaciones económicas, políticas, e institucionales/culturales del uso de las evaluaciones, sostienen que esto tiene más probabilidades de suceder en organizaciones con presión fiscal (p. 4), disponibilidad de recursos en la organización, tercerización en la producción de bienes y servicios (p. 5), mayoría de izquierda en el consejo local, alto nivel de competencia política, tendencia a seguir el modelo administrativo de moda (p. 7), grandes competencias y capacidades administrativas, rutinas de cambio bien desarrolladas, y el régimen político (p. 21).

En la misma línea explicativa Saliterer y Korac (2013) dividen los factores explicativos del uso de la información entre factores ambientales, inter-organizacionales, intra-organizacionales, e individuales (p. 4). Entre los ambientales, ubican la situación financiera de la organización (incluyendo su presión fiscal), y el tamaño de la organización; entre los inter-organizacionales están una escasa proporción de corporaciones públicas con actividades determinadas por el mercado, el grado de comercialización de los servicios locales, y los acuerdos cooperativos para la entrega de los mismo; los intra-organizacionales son: una cultura organizacional racionalista, un estilo de implementación gerencial, y un bajo o nulo estilo de implementación orientado hacia la legitimación; finalmente, en relación con los individuales, se subraya las actitudes en favor de las reformas de la Nueva Gestión Pública.

Kroll (2015) identifica como factores que tienen un impacto positivo para el uso de la información los siguientes: la madurez del sistema de medición, cuyas dimensiones pueden ser la facilidad del acceso a los indicadores de desempeño, tener metas exigentes pero alcanzables, y la satisfacción de las necesidades de información; el involucramiento de los interesados, que puede medirse con el grado de apoyo a las medidas de desempleo por parte de ciudadanos o funcionarios elegidos, o la ayuda por parte de los ciudadanos para evaluador el desempeño; el liderazgo, que puede medirse con el apoyo y compromiso para alcanzar los resultados que tiene el liderazgo superior de la organización, o que lo más alto de

la jerarquía organizacional enfatice y se preocupe por el proceso de gerencia del rendimiento de la organización; la capacidad organizacional, cuyas dimensiones implican recursos de tiempo, dinero y personas; una cultura de la innovación; claridad acerca de las metas; rutinas y foros de aprendizaje; actitudes positivas hacia las mediciones del desempeño, interacciones en red con otros interesados de la intervención medida; apoyo político; y un ambiente social fragmentado o diverso (p. 30-31).

Sandison (2006) a parte de especificar los factores relacionados con la calidad del proceso de medición que inciden en su uso, también identifica factores relacionales en donde se subraya la habilidad del evaluador en establecer relaciones constructivas con los usuarios clave; el rol y la influencia que tiene la unidad que realiza la evaluación, que debe guardar su independencia al tiempo que fomenta una relación cercana con los usuarios; y las redes que tienen los usuarios y los evaluadores, con las cuales se puede fortalecer la influencia y la credibilidad de la evaluación (p. 23). También establece factores organizacionales como una cultura organizacional de aprendizaje y apertura al escrutinio y el cambio; una estructura en donde se establezca una unidad organizacional dedicada a la rendición de cuentas y al aprendizaje, y la cual tenga una relación cercana con los tomadores de decisiones; y una administración del conocimiento a partir de mecanismos administrativos (p. 27).

Balthasar (2009) divide los factores explicativos del uso de la información de evaluaciones en factores institucionales que afectan la distancia entre los evaluadores y los evaluados; factores ambientales de la evaluación; y características del proceso de la evaluación. En los factores institucionales ubica la acción de la unidad responsable de la evaluación como interface entre los evaluadores y los evaluados, el financiamiento suficiente para la evaluación; la activación de la evaluación de manera independiente de aquellos responsables del desempeño medido, la realización de una evaluación a partir de una licitación; la publicación y accesibilidad de los resultados de la evaluación, y la realización de la evaluación de manera externa a la organización encargada de la intervención evaluada (p. 7). En los factores ambientales ubica el tipo de información requerida, la rutina de la evaluación, el clima político, la disponibilidad de otras fuentes de información, las características de los tomadores de decisión, y la apertura de estos al cambio. Finalmente, con relación a las características del proceso de evaluación incluye el potencial de los resultados de la evaluación para ser implementados, la razón entre costos y beneficios, difusión de la evaluación, facilidad de comprensión de la evaluación, y alta calidad del proceso de evaluación (p. 6).

Appleton-Dyer et al. (2012) dividen su modelo explicativo entre atributos particulares de la evaluación, como el enfoque, rol del evaluador, habilidades del evaluador, calidad del proceso, credibilidad, oportunidad, comunicación, y sensibilidad; características generales de la asociación para la evaluación, como el propósito, el tamaño, la complejidad y el nivel de colaboración; características del funcionamiento de esta asociación como la claridad en el propósito, la apropiación del proceso de la evaluación, confianza, poder, comunicación, cultura de aprendizaje, liderazgo,

toma de decisiones, intercambio de información, adaptabilidad del proceso de evaluación; comportamiento de la asociación para la evaluación como la activa participación, el apoyo para la evaluación, y la disponibilidad para la evaluación; características de los individuos tales como el involucramiento en la asociación, el rol que tienen en la misma, la disponibilidad que tienen para la evaluación, y las actitudes y conocimientos existentes; también subrayan la importancia de algunos factores contextuales como el apoyo a la evaluación, los recursos destinados para la evaluación, los requerimientos en materia de rendición de cuentas, el contexto de política pública y de administración, y mercados competitivos.

Por último, cerrando los modelos multi-teóricos, Taylor (2011) explica el uso de la información sobre rendimiento a partir de la calidad del sistema de medición del desempeño, una cultura organizacional jerárquica y racional, el apoyo de los grupos de interés al sistema de medición del desempeño de la agencia implementadora, un ambiente político externo apropiado, la percepción de los funcionarios sobre el impacto de las mediciones sobre el desempeño, y factores externos individuales, técnicos, y culturales.

1.1.3.3. Principales posturas acerca de la metaevaluación y los estándares de calidad de la practica evaluativa.

En los últimos años de la década de 1970, los funcionarios de los programas evaluados, así como los financiadores de las evaluaciones se volvieron cada vez más abiertamente escépticos frente a gastar recursos escasos en estudios que no podían entender o hallar relevantes. Por esta razón, a los evaluadores se les pidió "rendir cuentas" a partir una serie de preguntas como: ¿Cómo evaluar a los evaluadores?; y ¿cómo debe ser evaluada una evaluación? Fue en este contexto, según Patton (2002, pp. 549-550), que los evaluadores profesionales iniciaron la discusión acerca de los estándares de la práctica profesional.

Muy de la mano con un consenso profesional acerca de las explicaciones racionalistas centradas en la calidad del proceso de evaluación, estos criterios o estándares permitirían, a partir de su cumplimiento o satisfacción, brindar la percepción de credibilidad necesaria para que la información fuera utilizada: "It all depends on criteria. Judging quality requires criteria. Credibility flows from those judgments. Quality and credibility are connected in that judgments of quality constitute the foundation for perceptions of credibility" (2002, p. 542). Considerando que los evaluadores tienen el desafío de tomar responsabilidad por el uso de sus evaluaciones, en dado caso de que este uso no se realice, es necesario mirar en dónde las practicas o procesos han sido inadecuados (2012, p. 5), es decir, observar en dónde existen desfases entre lo normativo y lo realizado.

A partir de estas preocupaciones se desarrolla en la literatura relativa a la evaluación, el concepto de metaevaluación (por Scriven a finales de los años sesenta del siglo XX). Este concepto denota la evaluación de la evaluación, la cual

se establece como una obligación profesional de los evaluadores (2011, p. 388). Una revisión de la literatura acerca de la metaevaluación evidencia la existencia de cuatro principales posturas desarrolladas.

En primer lugar, el británico Michael Scriven establece que la metaevaluación puede orientarse a tres principales objetivos: juzgar el mérito de una evaluación (2010, p. 2); verificar la calidad de la gestión para obtener una retroalimentación que permita actuar sobre el proceso evaluativo (2015, p. 29); o examinar todas las actividades y procedimientos, especialmente el proceso general de aprendizaje / formación del programa (2015, p. 30). Dentro de los criterios que estableció para la realización de la metaevaluación, están los siguientes:

1. Validez: Este es el criterio clave en la cuestión de la veracidad de los hallazgos. Está compuesto por dos componentes: i) determinar qué tipo de evaluación es requerida junto con las limitaciones contextuales. Se necesita saber qué tipo de evaluación fue requerida en principio sumativa o formativa; ii) probabilidad de verdad de la conclusión, este proceso involucra dos dimensiones, 1. Cobertura: en donde se vuelve importante tener una lista de verificación específica del área, una descripción correcta del evaluando y de las sub-evaluaciones de proceso, resultados, costos, alternativas y generalización. 2. Corrección: significa que se cumplen los estándares científicos relevantes; se verifica la idoneidad de la evidencia y las inferencias que se proporcionan para respaldar las conclusiones evaluativas propuestas en la evaluación objetivo (2010, p. 2 -3)
2. Credibilidad: Combinación de bajo nivel de sesgo probable, por ejemplo, de conflicto de intereses con un alto nivel de experiencia. (2010, p. 4)
3. Claridad: combinación de comprensibilidad para el cliente, audiencias, tomadores de decisiones, personal y resultados concisos, ambos factores reducen el esfuerzo y la frecuencia de errores de interpretación, y pueden mejorar la aceptación e implementación (2010, p. 5)
4. Propiedad: requiere ética, legalidad y adecuación cultural, en la medida en que se puedan combinar: esto debe incluir la consideración del respeto de las obligaciones contractuales (por ejemplo, plazos), la privacidad, el consentimiento informado y evitar la explotación de grupos de clase social, género, edad religiosa, étnica u orientación sexual. (2010, p. 5)
5. Costo-utilidad: significa "ser económico" en términos de sentido común, pero también cubre el análisis de costos y beneficios que incluye el contexto y las ganancias y pérdidas de capital ambiental, personal y social. Esto normalmente requiere: (i) que la evaluación al menos incluyera una cuidadosa estimación de costo-factibilidad y (ii) al menos una estimación de rentabilidad comparativa (2010, p. 5)
6. Generalización: no es un criterio de mérito necesario o definitorio, por lo que no hay una "barrera" que deba superarse en esta dimensión, pero es un factor de ganancia adicional, por lo que los diseñadores y profesionales de la evaluación deben buscar tenerlo presente. Esto tiene dos facetas de particular importancia: (i) utilidad: mérito de este diseño de evaluación y en el procedimiento para su uso; (ii) solidez, es decir, la medida en que los

> resultados de la evaluación son inmunes a un grado de cambios en el contexto del programa o variaciones del programa del tipo habitual relativamente menor. (2010, p. 5 - 6)

Teniendo en cuenta estos criterios, la perspectiva de Scriven es posible agruparla fácilmente dentro de las perspectivas que privilegian la calidad de la información resultante de la evaluación, como estándares para evaluar las evaluaciones. Particularmente, los criterios de validez y credibilidad se orientan a estándares científicos que deben ser satisfechos por el proceso evaluativo en aras de tener información robusta. A pesar de que se tienen en cuenta elementos como la utilidad o la claridad que tiene la evaluación para los interesados, no se toca ningún elemento relativo a la relación entre los clientes y los evaluadores.

Una segunda perspectiva importante de ser resaltada en metaevaluación es la de Thomas Schwandt, el cual identificó como posibles objetivos de la metaevaluación la realización de seguimientos y mejoras a la calidad del trabajo de evaluación (1992, p. 95); examinar los procedimientos y hallazgos de una evaluación con el fin de juzgar la medida en que los procedimientos se ajustan al criterio de confiabilidad y el grado en que los hallazgos cumplen el criterio de confiabilidad (1992, p. 96); y examinar los datos disponibles, la elección, descripción, y uso de métodos de evaluación (1992, p. 96).

En un texto escrito en conjunto entre Schwandt y Edward Halpern establecen que la metaevaluación debe servir para el mejoramiento de la calidad de las evaluaciones. En este libro establecen como principales criterios la comprobación de la pertinencia de la metodología de la evaluación y, por otro lado, la verificación de si la información producida por la evaluación es o no correcta, para lo que se deberían abordar criterios como relevancia, fiabilidad, suficiencia, representatividad y oportunidad (1988). La perspectiva de Schwandt y Halpern también puede ser clasificada como integrante de las que privilegian la calidad de la información y el rigor metodológico como criterios de metaevaluación.

A inicios del siglo XXI los finlandeses Uusikylä y Virtanen escribieron un artículo en donde desarrollan un enfoque de la metaevaluación orientado al aprendizaje, en donde este ejercicio debe enmarcarse en un diálogo abierto entre las distintas partes interesadas del proceso de evaluación a evaluar (2000). A partir de un estudio de 15 evaluaciones de mediano término sobre programas implementados en Finlandia, concluyen que un análisis de evaluaciones particulares sirve tanto para el aprendizaje de política como el organizacional, entendiendo el aprendizaje en política como los mecanismos de diseminación del conocimiento en las distintas arenas políticas, y el organizacional como la investigación evaluativa continua y reflexiva (p.1).

Dentro de los criterios de metaevaluación que son utilizados por parte de Uusikylä y Virtanen fueron los lineamientos de la evaluación, la selección de los evaluadores, la planeación e implementación del proceso de evaluación, la calidad de los reportes de evaluación, la interacción entre los actores involucrados, y la valoración de la experiencia de la evaluación como un todo. Es importante desatacar que los autores hacen particular énfasis en el diálogo y la transparencia de la evaluación como los principios claves para valorar el mérito y la bondad de los procesos evaluativos.

La referencia más importante (y ampliamente reconocida) en cuanto a metaevaluación es Daniel Stufflebeam, esto es así, porque este autor además de haber diseñado una lista de criterios que ha sido ampliamente aceptado dentro de la disciplina evaluativa, ha venido desarrollando y consolidando sus primeras posturas en compañía de otros autores relevantes en el campo. Stufflebeam ha manifestado que la metaevaluación puede orientarse a tres tipos de objetivos: delinear, obtener y aplicar información descriptiva e información de juicio sobre la utilidad, viabilidad, propiedad y precisión de una evaluación. (2001, p. 183); guiar la evaluación e informar públicamente su fortalezas y debilidades. (2001, p. 183); ayudar a los evaluadores a planificar, realizar, mejorar, interpretar e informar sobre sus estudios de evaluación. (2001, p. 185).

Stufflebeam construyó en 1999 una lista de chequeo denominada *Program evaluations metaevaluation checklist (based on The Program Evaluation Standars)*, la cual tenía como propósito realizar una metaevaluación sumativa, y fue organizada en correspondencia con los Joint Committee Program Evaluation Standars. En este documento, se incluyen los siguientes estándares:

- Utilidad: identificación de interesados, credibilidad del evaluador, alcance y selección de la información, identificación de valores, claridad del reporte, oportunidad y difusión del reporte e impacto de la evaluación.
- Factibilidad: procedimientos prácticos, viabilidad política y costo-efectividad.
- Conformidad: Orientación al servicio, acuerdos formales, derechos de las personas, interacciones humanas, valoración completa y justa, divulgación de hallazgos, conflicto de intereses, responsabilidad fiscal)
- Exactitud: Documentación, análisis de contexto, descripción de propósitos y procedimientos, fuentes de información, validez de información, confiabilidad de información, información sistemática, análisis de información cuantitativa y cualitativa, conclusiones justificadas, reportes imparciales, y metaevaluación)

Posteriormente, Sufflebeam junto con Goodyear, Marquart y Johnson, construyeron una segunda lista de chequeo para metaevaluación (2005). Este documento, en contraste con el anterior, tiene una finalidad de establecer lineamientos para metaevaluaciones tanto formativas como sumativas. Los autores claramente establecen que esta lista es una revisión de su anterior versión (p. 1). Dentro de los objetivos y usos que se le pueden dar a esta nueva lista de chequeo se encuentran: evaluar si la evaluación objetivo cumplió con los objetivos propuestos (mejorar productos, personal, programas, organizaciones, gobiernos, consumidores y el

interés público; contribuir a la toma de decisiones informada y un cambio más inteligente); y contribuir al cambio necesario para empoderar a todos los interesados mediante la recopilación de datos de ellos y su participación en el proceso de evaluación y generando nuevos conocimientos. (p. 4). Dentro de los criterios de metaevaluación aclarados en este documento se encuentran:

1. Investigación sistemática: conocimiento de estándares técnicos para la aplicación de métodos cuantitativos y cualitativos para obtener e informar de forma precisa y creíble información (p. 9); Involucrar al cliente en la exploración de las deficiencias y fortalezas de una gama razonable de posibles preguntas de evaluación. (p. 10); Involucrar a una gama adecuada de partes interesadas en la exploración de las deficiencias y fortalezas de una gama razonable de posibles preguntas de evaluación; Informar con suficiente detalle el enfoque empleado y los métodos para permitir que otros comprendan, interpreten y hagan una crítica a la evaluación, al proceso y los hallazgos (p. 10).
1. Competencia: Poseer la educación habilidades, destrezas y experiencia apropiada; nivel suficiente de competencia cultural para asegurar el reconocimiento, exacta interpretación, y respeto por diversidad; Práctica dentro de los límites de su formación profesional (p. 14).
2. Integridad y honestidad: Tomar la iniciativa de negociar todos los aspectos de la evaluación con los clientes y las partes interesadas representativas; Negociar honestamente con clientes y representante interesados relativa a las limitaciones de los métodos seleccionados; Negociar honestamente con clientes y representantes interesados sobre el alcance de resultados que probablemente sean adquiridos (2005, p. 19); Al proceder con la evaluación, informe claramente cualquier conflicto de intereses que hayan tenido y cómo se abordaron (p. 20).
3. Respeto por las personas: Desarrollar una comprensión de la contextual de la evaluación elementos, incluyendo, como apropiado, geográfico ubicación, tiempo, político y clima social, condiciones económicas, y otras dinámicas relevantes en progresar al mismo tiempo; cumplir con la ética profesional, los estándares y las regulaciones actuales con respecto al consentimiento informado de los participantes de la evaluación; Realizar la evaluación y comunicar los resultados de una manera que respete claramente la dignidad y autoestima de cada parte interesada; Asegurar que los participantes de la evaluación tuvieran pleno conocimiento y la oportunidad de obtener cualquier beneficio de la evaluación (2005, p. 26 -27-28).
4. Responsabilidades generales y bienestar público: Presentar planes e informes de evaluación que incluyan, según corresponda, perspectivas e intereses relevantes de toda la gama de interesados; Considere no solo las operaciones y los resultados inmediatos del evaluando, sino también sus supuestos generales, implicaciones y posibles efectos secundarios; Informar de forma clara y sencilla para que los clientes y otros las partes interesadas podrían fácilmente entender la evaluación proceso y resultados; Mantener un

> equilibrio adecuado entre satisfacer las necesidades del cliente y otras necesidades (2005, p. 32 - 33).

Los estándares y criterios de Stufflebeam se plasmaron institucionalmente en el Joint Committee on Standards for Educational Evaluation (Alianza EEUU-Canadá). Este comité fue creado en 1975 y es una coalición de asociaciones profesionales desarrollada para desarrollar estándares de evaluación y mejorar la cualificación de la estandarización de la evaluación.

Para realizar su misión este comité revisa y actualiza sus estándares cada cinco años. Stufflebeam fue presidente de este comité (de 1975 a 1988), miembro por más de 13 años, y fue el autor principal de la versión original de los estándares de evaluación propuestos por el comité, y coautor de las posteriores reformulaciones. La siguiente es la lista actual de estándares/criterios de evaluación[28] del Joint Committe on Standars for Educational Evaluation (evaluación para programas):

- *Estándares de utilidad*: Credibilidad del evaluador; Atención a las partes interesadas; Negociación de propósitos de la evaluación; Valores explícitos; Información relevante; Procesos y productos significativos; Comunicaciones y reportes oportunos y apropiados; Interés por las consecuencias y la influencia de la evaluación.
- *Estándares de factibilidad*: Estrategias efectivas de administración; Procesos prácticos y sensibles al modo de implementación del programa evaluado; Viabilidad contextual; Uso de recursos eficientes y efectivos.
- *Estándares de conformidad:* Orientación sensible e inclusiva; Negociación de acuerdos formales; Respeto a los derechos humanos; Claridad y equidad; Transparencia y divulgación; Conflictos de interés explícitos; Responsabilidad fiscal.
- *Estándares de precisión*: Conclusiones y decisiones justificadas; Información válida; Información confiable; Descripciones explicitas del programa y del contexto; Gestión de la información; Diseño y análisis apropiados; Razonamiento de la evaluación explícito; Comunicación y reporte.
- *Estándares de rendición de cuentas*: Documentación del proceso de evaluación; Metaevaluación interna; Metaevaluación externa.

Existen otras dos institucionalizaciones de estándares de calidad para la práctica evaluativa: los de la OCDE (2010) y los de la ONU (2005), las cuales también establecen lineamientos de buenas prácticas y brindan herramientas para poder evaluar los procesos evaluativos. Estas dos institucionalizaciones están

[28] Consultado en: https://evaluationstandards.org/program/ , También puede encontrarse en Yarbrough (2011)

fuertemente inspiradas en los estándares del Joint Committee, lo cual se evidencia en que comparten un número considerable de estándares:

Estándares de Calidad para la Evaluación del Desarrollo de la OCDE (OCDE, 2010)

- *Estándares generales*: Proceso evaluativo libre y abierto; Ética de la evaluación; Enfoque asociativo; Coordinación y alineamiento; Desarrollo de capacidad; Control de calidad.
- *Estándares de finalidad, planificación y diseño*: Fundamentación y finalidad de la evaluación; Objetivos específicos de la evaluación; Objeto y alcance de la evaluación; Evaluabilidad de la intervención; Participación de los actores relevantes; Consideración sistemática de la evaluación conjunta; Cuestiones a las que pretende responder la evaluación; Selección y aplicación de criterios relevantes de evaluación; Selección del enfoque y de la metodología; Recursos de la evaluación; Estructuras de gobierno y gestión; Documentación relativa a la finalidad y las expectativas de la evaluación.
- *Estándares de ejecución de la evaluación y presentación de los resultados*: Equipo de la evaluación; Independencia de los evaluadores frente a los interesados; Consulta y protección de los implicados; Realización de la evaluación dentro del plazo y el presupuesto asignados; Informes de evaluación; Claridad y representatividad del resumen; Contexto de la intervención; Lógica de la intervención; Validez y fiabilidad de las fuentes de información; Explicación de la metodología utilizada en el informe de evaluación; Claridad del análisis; Respuestas a las preguntas de evaluación; Mención de los cambios y de las limitaciones de la evaluación; Mención de las discrepancias entre los miembros del equipo; Inclusión de las observaciones de los actores implicados.
- *Estándares de seguimiento, utilización y aprendizajes*: Oportunidad, pertinencia y utilización de la evaluación; Respuesta sistemática a las recomendaciones y seguimiento de las mismas; Difusión de la evaluación.

Estándares para la Evaluación en el sistema de Naciones Unidas (UNEG, 2005).

- *Estándares de marco institucional*: marco institucional adecuado para la gestión de la evaluación; Desarrollo de una política de evaluación y su actualización; Revisión y aprobación de planes de evaluación por parte de directivos; mecanismos de seguimiento y difusión adecuados;
- *Estándares de gestión de la evaluación*: Líder de la evaluación se asegura de cumplir los estándares de evaluación; Líder de la evaluación es responsable por preparar las directrices de la evaluación; Líder de la evaluación asegura que esta sea dinámica y adaptable a nuevos desarrollos y necesidades.

- *Estándares de competencias*: Evaluadores poseen competencias en evaluación; Evaluadores poseen educación relevante; Evaluadores tienen experiencia relevante; Evaluadores poseen conocimiento técnico específico, así como habilidades técnicas y gerenciales específicas.
- *Estándares Éticos*: Evaluadores son sensibles a creencias y costumbres y actúan con integridad con todos los interesados de la intervención; Respeto caracteriza todas las relaciones de los evaluadores; Evaluadores protegen el anonimato y confidencialidad de los informantes; Los evaluadores son responsables de su desempeño y de sus productos.
- *Estándares de diseño*: Diseño asegura información oportuna, válida, confiable y relevante; Términos de referencia determinan el propósito y describe procesos y productos de la evaluación; Propósito y contexto de la evaluación explícitos; Descripción clara del objeto a ser evaluado; Objetivos de la evaluación son realistas y alcanzables; Declaración puntual de los criterios de evaluación; rigor metodológico; valoración de costo-efectividad; inclusión del enfoque de derechos humanos del Sistema de Naciones Unidas.
- *Estándares de proceso*: Relación de confianza y respeto entre evaluadores y comisionado de la evaluación; Consulta de las partes interesadas para todo el proceso de evaluación; Revisión de pares y de expertos externos.
- *Estándares de la selección del equipo evaluador*: Evaluaciones conducidas por equipos bien cualificados; Composición del equipo debe tener balance de género y geográfico.
- *Estándares de implementación*: Evaluación implementada de manera ética y profesional.
- *Estándares de reportes*: Estructuración lógica del reporte de evaluación; La página de apertura provee información básica clave; Resumen ejecutivo; Descripción del modelo lógico o cadena de resultados de la intervención; Descripción de los roles y contribuciones de los interesados en la intervención; Descripción del propósito y contexto de la evaluación; Reporte explica los criterios de evaluación; Reporte explica los objetivos y el alcance de la evaluación; Reporte indica el grado en el que se incorporó el enfoque de derechos humanos y de género; la metodología es descrita de modo transparente incluyendo sus límites; Descripción de la participación de los interesados; Descripción de mecanismos de protección de los participantes; Insumos, productos y resultados e impactos deben medirse tanto como sea posible; análisis incluye discusión sobre la contribución de los interesados a los resultados; Identificación de éxitos y dificultades de la intervención; Conclusiones son sustentadas por los hallazgos; Recomendaciones son relevantes, realistas y se basan en la evidencia y en el análisis; Lecciones deben ser generalizadas más allá de la intervención evaluada; Anexos deben ser completos y relevantes.
- *Estándares de seguimiento*: Respuesta explícita de las autoridades acerca de la respuesta y la gestión de las recomendaciones de la evaluación.

A pesar de que esta es una caracterización bastante sintética de estas listas, en la que únicamente se denotan a grandes rasgos los estándares, sin profundizar en la

conceptualización de todos y cada uno de estos, se puede observar que las tres listas de estándares (institucionalizados) consideran central la utilización de la información producida por la evaluación y, en general, abordan los mismos temas, lo cual evidencia un considerable nivel de consenso académico alrededor de lo que puede denominarse como la calidad de la evaluación.

Sin embargo, la primera lista de estándares (Joint Committee) toca de manera muy superficial temas como las características que deben tener los informes y los reportes de evaluación, la metodología de la evaluación y los criterios para su selección y, si bien se establece un necesario interés por las consecuencias y la influencia de la evaluación, este estándar se queda corto frente al objetivo del uso de la evaluación y el seguimiento que se le debe realizar.

Por su parte, la tercera lista aborda varios temas que son específicos a las intervenciones que cuentan con cooperación internacional del Sistema de Naciones Unidas, como la relación de confianza y respeto entre evaluadores y comisionado de la evaluación, la inclusión del enfoque de derechos humanos y de género dentro de la evaluación, la descripción de las contribuciones de los interesados en los resultados de la intervención. Además de algunas redundancias como los estándares de competencias con los de selección del equipo evaluador.

La segunda lista de estándares, correspondiente a la de la OCDE, ayuda a entender la evaluación de políticas como un proceso, con determinados estándares para cada una de sus etapas. No sufre de redundancia o de irrelevancia en sus temas frente a la política de evaluación implementada en Colombia. Refleja de mejor manera un consenso académico racionalista en las variables que maneja y en su proceso de construcción[29].

Teniendo en consideración las guías metodológicas de la política pública de evaluación en Colombia, se observa una fuerte inspiración en las metodologías de la OCDE[30], las cuales pueden establecerse como la principal institucionalización

[29] Al respecto, la Red de Evaluación del Comité de Ayuda al Desarrollo de la OCDE puntualizó: “En su versión provisional, los Estándares se aprobaron en 2006 para un período de prueba de tres años y se han revisado a la luz de la experiencia. Toda una serie de socios para el desarrollo han contribuido a este proceso, incluidos donantes y países asociados. La aportación inicial se fraguó en un seminario celebrado en 2005 en Nueva Delhi. Con una encuesta en 2008 sobre el uso de los Estándares y un seminario celebrado en 2009 en Auckland, más los comentarios recibidos por los miembros de la Red del DAC de Evaluación del Desarrollo se consiguió perfeccionar y ultimar el texto, con el apoyo de la Secretaría de la OCDE. La mencionada Red aprobó la versión definitiva el 8 de enero de 2010, que fue ratificada por el CAD el 1 de febrero de ese mismo año” (OCDE, 2010, p. 7).

[30] Esto se puede evidenciar con el considerable número de referencias que se realizan en las guías del DNP a las guías publicadas por la OCDE, y por la coincidencia entre los estándares y los lineamientos con los que se han construido estas guías, tal como se evidenciará más adelante.

occidental del consenso racionalista entorno a la evaluación de políticas públicas. Por estas razones, la operacionalización de la primera variable independiente de la hipótesis de la presente investigación (satisfacción de los estándares profesionales de calidad), se realizará a partir de la teorización normativa realizada por la OCDE en su lista de estándares para la evaluación.

Para construir un puente entre esta teorización normativa y el marco de análisis de política pública de Ostrom, es pertinente proceder a ubicar cada uno de los estándares de calidad dentro de las categorías conceptuales del IAD, esto permitirá en los capítulos posteriores realizar un contraste entre lo normativo y lo implementado al interior de cada una de las variables de Ostrom, así como clarificar los estándares específicos que al funcionar como reglas en las situaciones de acción, funcionan como variables explicativas del comportamiento de los actores que buscan satisfacerlos. Todos los estándares considerados pueden ser conceptualmente organizados dentro de uno o varios de los tipos de reglas que determinan las variables de la situación de acción:

Tabla empalme variables IAD-Estándares profesionales de calidad

Tipo de regla (Ostrom)	Estándar (OCDE)	Conceptualización del estándar
Arena de acción de diseño de la evaluación		
Reglas de frontera	Enfoque asociativo	"El concepto de asociación lleva implícita la idea de proceso incluyente en el que participan diversos actores interesados, como los poderes ejecutivo y legislativo, la sociedad civil, los supuestos beneficiarios y los socios internacionales" (OCDE, 2010, p. 7).
	Participación de los actores relevantes	"Los actores implicados relevantes deben intervenir en las primeras fases del proceso y debe dárseles la oportunidad de contribuir a su diseño, incluyendo la identificación de los asuntos que han de abordarse y las preguntas a las que la evaluación ha de responder" (p. 9).
	Consideración de evaluación conjunta	"A fin de contribuir a la armonización, la alineación y una eficaz división del trabajo, los organismos donantes y los países socios deben sopesar de forma sistemática la posibilidad de efectuar una evaluación conjunta, en régimen de colaboración entre más de un organismo o país (…) Las evaluaciones conjuntas abordan cuestiones de interés común para los socios, así como materias específicas que afectan individualmente a cada uno de ellos" (p. 9).
Reglas de posición	Independencia de los evaluadores	"Los evaluadores deben ser independientes respecto de la intervención para el desarrollo de que se trate, incluido de sus funciones de formulación de políticas, operativas y de gestión, así como sus beneficiarios previstos. Los posibles conflictos de interés se tratan abierta y francamente. El equipo evaluador debe trabajar con libertad y sin injerencias. Debe asegurársele cooperación y acceso a toda la información pertinente" (p. 11).
	Ética de la evaluación	"Los evaluadores respetan las principales normas y códigos de conducta profesionales y éticos que les atañen. La evaluación se lleva a cabo con integridad y honradez. Los directores y gestores de la evaluación y los evaluadores deben respetar los derechos humanos y las diferencias

		culturales y en las costumbres, creencias y prácticas religiosas de todos los actores implicados." (p. 6).
	Estructuras de gobierno y gestión	"Las estructuras de gobierno y gestión se deben diseñar para adaptarse al contexto, propósito, alcance y objetivos de la evaluación. La estructura de gobierno debe salvaguardar la credibilidad, la transparencia y el carácter integrador de la evaluación. La estructura de gestión debe organizar el proceso y ocuparse de su gestión cotidiana. En función de las dimensiones y complejidad de la evaluación, estas funciones pueden combinarse o ejercerse por separado." (p. 10).
Reglas de elección	Objeto y alcance de la evaluación	"La intervención para el desarrollo evaluada (el objeto de la evaluación) debe definirse claramente, incluida la descripción de la lógica o teoría de la intervención. El alcance de la evaluación definirá el plazo, los fondos desembolsados, la zona geográfica, los grupos destinatarios, la configuración organizativa, los mecanismos de aplicación, el contexto estratégico e institucional y otras dimensiones que debe abarcar el ejercicio. Se pondrán de manifiesto las discrepancias entre la aplicación prevista y la efectiva ejecución de la intervención" (p. 8).
	Evaluabilidad de la intervención	"La viabilidad de la evaluación debe ser analizada. Más en concreto, debe determinarse si se ha definido adecuadamente la intervención, si sus resultados son verificables y si la evaluación constituye la mejor respuesta a las cuestiones planteadas por los actores implicados relevantes o por los responsables de las políticas" (p. 9).
	Selección y aplicación de criterios de evaluación	"En la evaluación se aplican los criterios adoptados por el CAD referentes a la evaluación de la ayuda al desarrollo: pertinencia, eficiencia, eficacia, impacto y sostenibilidad. La aplicación de estos y otros criterios adicionales depende de las preguntas evaluativas y los objetivos de la evaluación. Si se descarta uno de ellos y/o se añaden otros criterios, deberá explicarse en el informe de evaluación. Todos los criterios aplicados deberán definirse en términos inequívocos" (p. 9).
	Cuestiones para responder	"Los objetivos de la evaluación se plasman en una serie de preguntas evaluativas específicas y relevantes, que deben decidirse en una fase temprana del proceso e inspirar la elaboración de la metodología." (p. 9).
	Selección del enfoque y de la metodología	"El propósito, alcance y preguntas evaluativas determinan el enfoque y la metodología apropiados para cada caso. Cabe emplear un informe inicial para explicar la selección del enfoque elegido (…) La metodología se elabora de acuerdo con el enfoque elegido y debe incluir la especificación y justificación del diseño de la evaluación y las técnicas de recogida y análisis de datos. La metodología seleccionada debe responder a las preguntas de la evaluación utilizando evidencias creíbles. Se debe distinguir entre los diferentes niveles de resultados (lógica de intervención que contenga una jerarquía entre objetivos y medios, indicando los insumos, los productos, el efecto directo y el impacto). Los indicadores para medir el logro de los objetivos se validan de conformidad con criterios generalmente aceptados, por ejemplo, el que responde a las siglas SMART en inglés (Específicos, Medibles, Alcanzables, Realistas y Temporalmente determinados). Deben presentarse los datos desglosados para clarificar cualesquiera diferencias entre sexos y distintos grupos de pobres ¿pobres? incluidos los grupos excluidos (p. 10).
	Coordinación y alineamiento	"Para mejorar la coordinación y reforzar los sistemas nacionales, el proceso de evaluación ha de tener en cuenta los planes, actividades y políticas nacionales y locales en la materia." (p. 7).

	Fundamentación y finalidad de la evaluación	"La fundamentación, finalidad y utilización prevista de la evaluación deben indicarse claramente, mencionándose por qué se efectúa la evaluación en ese momento concreto, para quién se lleva a cabo y cómo se emplearán sus resultados a efectos de aprendizaje y rendición de cuentas" (p. 8).
	objetivos específicos de la evaluación	"Los objetivos específicos de la evaluación deben establecer claramente el propósito de esta. (…) determinar resultados (productos, efectos e impactos) y analizar la eficacia, eficiencia, pertinencia y sostenibilidad de una intervención específica para el desarrollo (…) ofrecer hallazgos, conclusiones y recomendaciones respecto a una intervención concreta a fin de extraer enseñanzas para futuras actividades en diseño y ejecución." (p.- 8).
	Realizar la evaluación dentro del plazo y el presupuesto asignados	"La evaluación debe realizarse y sus resultados facilitarse a quienes la encomendaron en la forma oportuna y que permita alcanzar los objetivos previstos. La evaluación debe llevarse a cabo de forma eficiente y dentro del presupuesto asignado. Las modificaciones de las condiciones y circunstancias han de ponerse de manifiesto, y se explicarán, tratarán y consensuarán entre las partes los cambios imprevistos de plazo y presupuesto." (p. 11).
Reglas de información	Documentación relativa a la finalidad y las expectativas	"La fase de planificación y diseño culmina en la redacción de un documento claro y completo que suele denominarse "Términos de Referencia" (TdR) y en el que se presentan la finalidad, el alcance y los objetivos de la evaluación; la metodología que se aplicará; los recursos y el plazo asignados; los requisitos de presentación de informes y cualesquiera otras expectativas referidas al proceso y productos de la evaluación. El documento debe ser consensuado por el gestor o gestores de la evaluación y los que vayan a llevarla a cabo. Este documento también puede recibir otras denominaciones, como 'alcance del trabajo' o 'mandato de evaluación'." (p. 10).
Reglas de agregación	Control de calidad	"El control de calidad se ejerce a lo largo de todo el proceso de evaluación. En función del alcance y complejidad de este último, el control de calidad debe canalizarse a través de mecanismos internos y/o externos, por ejemplo, revisión inter pares, grupos consultivos o grupos de referencia" (p. 7).
Reglas de alcance	Documentación relativa a la finalidad y las expectativas	"La fase de planificación y diseño culmina en la redacción de un documento claro y completo que suele denominarse "Términos de Referencia" (TdR) y en el que se presentan la finalidad, el alcance y los objetivos de la evaluación; la metodología que se aplicará; los recursos y el plazo asignados; los requisitos de presentación de informes y cualesquiera otras expectativas referidas al proceso y productos de la evaluación. El documento debe ser consensuado por el gestor o gestores de la evaluación y los que vayan a llevarla a cabo. Este documento también puede recibir otras denominaciones, como 'alcance del trabajo' o 'mandato de evaluación'." (p. 10).
Reglas de pago	No existen estándares que puedan considerarse reglas de pago	No existen estándares que puedan considerarse reglas de pago.
Arena de acción de ejecución de la evaluación		
Reglas de frontera	Participación de actores relevantes	"Los actores implicados relevantes deben intervenir en las primeras fases del proceso y debe dárseles la oportunidad de contribuir a su diseño,

		incluyendo la identificación de los asuntos que han de abordarse y las preguntas a las que la evaluación ha de responder" (p. 9).
	Recursos	"Los recursos disponibles para la evaluación deben ser suficientes en cuanto a fondos, personal y habilidades técnicas, de manera que los objetivos de esta puedan cumplirse de forma eficaz." (p. 10).
	Equipo de la evaluación	"Debe seguirse un procedimiento de contratación abierto y transparente para la selección del equipo evaluador. Los integrantes del equipo deben dominar una combinación de habilidades en materia de evaluación y conocimientos temáticos. No debe olvidarse el equilibrio entre géneros; en el equipo deben incluirse profesionales de los países o regiones socios." (p. 11).
Reglas de posición	Independencia de los evaluadores	"Los evaluadores deben ser independientes respecto de la intervención para el desarrollo de que se trate, incluido de sus funciones de formulación de políticas, operativas y de gestión, así como sus beneficiarios previstos. Los posibles conflictos de interés se tratan abierta y francamente. El equipo evaluador debe trabajar con libertad y sin injerencias. Debe asegurársele cooperación y acceso a toda la información pertinente" (p. 11).
Reglas de elección	Consulta y protección de implicados	"Se debe consultar a todo el conjunto de actores relevantes —incluidos socios y donantes— durante el proceso de evaluación, brindándoles la oportunidad de contribuir. Se han de especificar los criterios para identificar y seleccionar a los interesados (…) Se debe proteger los derechos y el bienestar de los participantes en la evaluación. Cuando así se solicite o resulte necesario, se protegerá la confidencialidad y el anonimato de los informantes" (p. 11).
	Explicación de la metodología utilizada	"En el informe se debe describir y explicar la metodología evaluativa y su aplicación. Para analizar efectos directos e impactos, se explicará la atribución y/o contribución de los resultados. En el informe se reconocerán las posibles restricciones halladas y su efecto sobre la evaluación, incluyendo sobre su independencia e imparcialidad. Se han de detallarlas técnicas empleadas para la recogida y análisis de los datos. Se justificarán las opciones y se explicarán las limitaciones y deficiencias." (p. 13).
	Claridad del análisis	"En el informe de evaluación se deben presentar por separado los hallazgos, conclusiones, recomendaciones y lecciones extraídas, con una distinción clara y lógica entre ellos. Los hallazgos deben derivarse lógicamente del análisis de los datos, mostrándose una línea clara de razonamiento que sustente las conclusiones. Las conclusiones deben justificarse mediante los hallazgos y el análisis. Las recomendaciones y las lecciones que quepa extraer se deducirán lógicamente de las conclusiones." (p. 13).
	Lógica de la intervención	"En el informe de evaluación se describe y valora la lógica o teoría que sustenta la intervención, incluidos las suposiciones y factores subyacentes que pueden afectar al éxito de esta." (p. 12).
	Claridad y representatividad del resumen	"El informe escrito debe ir encabezado por un resumen ejecutivo, que ha de ofrecer una panorámica de su contenido y en la que se expondrán los principales hallazgos, conclusiones, recomendaciones y las lecciones generales que, en su caso, quepa extraer" (p. 12).

	Respuestas a las preguntas evaluativas	"En el informe de evaluación se deben responder las cuestiones detalladas en los TdR a efectos de evaluación y, si no fuera posible, debe explicarse por qué." (p. 13).
	Validez y fiabilidad de las fuentes de información	"En el informe de evaluación se describen las fuentes de información empleadas (documentos, informantes, datos administrativos, bibliografía, etc.) con el detalle suficiente para que pueda valorarse su idoneidad. Se debe explicar también la selección de estudios de caso o muestras, señalándose las limitaciones referentes a la representatividad de estas últimas. En la evaluación han de validarse de manera cruzada las fuentes de información y analizarse críticamente la validez y fiabilidad de los datos." (p. 12).
	Mención de discrepancias entre los miembros del equipo	"Los miembros del equipo de evaluación deben disponer de la posibilidad de discrepar de juicios y recomendaciones concretos con los que estén en desacuerdo. Cualesquiera diferencias de opinión en el seno del equipo deben mencionarse en el informe." (p. 13).
	Mención de los cambios y limitaciones de la evaluación	"En el informe de evaluación se expondrán las limitaciones que se hayan suscitado en relación con el proceso, metodología y datos, y se analizará la validez y fiabilidad del ejercicio. Se debe mencionar cualquier obstáculo que haya surgido para efectuar un proceso evaluativo libre y abierto y que pueda haber influido en los hallazgos. Se indicarán las discrepancias entre el proceso de aplicación y los frutos de este, según las previsiones iniciales y los resultados efectivos." (p. 13).
	Contexto de la intervención	"En el informe se debe describir el contexto de la intervención para el desarrollo, incluidos: (…) el contexto político, las políticas, objetivos y estrategias de la agencia de desarrollo y delos socios; (...) el contexto de desarrollo, incluidos los factores socioeconómicos, políticos y culturales; (...) el contexto institucional y la participación de los actores implicados. La evaluación debe permitir determinar y valorar la influencia del contexto en los resultados de la intervención para el desarrollo." (p. 12).
	Respuesta sistemática a las recomendaciones y seguimiento de estas	"Las personas/entidades destinatarias de las recomendaciones deben responder de forma sistemática a las mismas y adoptar las medidas procedentes. Esto incluye una respuesta y un seguimiento formales por parte de la dirección. Las medidas acordadas han de ser supervisadas para garantizar que se rinde cuentas de su aplicación." (p. 15).
Reglas de información	Difusión	"Los resultados de la evaluación se presentan en un formato accesible y se difunden de forma sistemática, tanto interna como externamente, para extraer lecciones aprendidas, realizar acciones de seguimiento y garantizar la transparencia. A la luz de las enseñanzas surgidas del ejercicio, se señalan otros posibles interesados dentro de la amplia comunidad del desarrollo y se entabla contacto con ellos de forma que se saque el máximo partido" (p. 15).
	Informe de la evaluación	"El informe de la evaluación ha de ser fácilmente comprensible por sus destinatarios, y su modalidad será la adecuada a la vista de la finalidad del ejercicio" (p. 11).
Reglas de agregación	Control de calidad	"El control de calidad se ejerce a lo largo de todo el proceso de evaluación. En función del alcance y complejidad de este último, el control de calidad debe canalizarse a través de mecanismos internos y/o externos, por ejemplo, revisión inter pares, grupos consultivos o grupos de referencia" (p. 7).

	Proceso evaluativo libre y abierto	"Para una mayor credibilidad, el proceso de evaluación ha de ser transparente e independiente de la gestión programática o la elaboración de políticas en la materia" (p. 6).
	Inclusión de observaciones de implicados	"Se debe ofrecer a los actores implicados relevantes la posibilidad de formular observaciones al borrador de informe." (p. 14).
Reglas de alcance	Oportunidad, pertinencia y utilización de la evaluación	"La evaluación debe diseñarse, dirigirse y presentar su informe con vistas a satisfacer las necesidades de los usuarios previstos. Las conclusiones, recomendaciones y lecciones se formulan de forma clara, pertinente, focalizada y orientada para su puesta en práctica, de tal modo que la evaluación pueda utilizarse para lograr los objetivos de aprendizaje y rendición de cuentas previstos. La evaluación se entrega puntualmente para permitir un empleo óptimo de sus resultados. Se garantiza una difusión sistemática y la conservación y gestión del informe de manera que todos los socios para el desarrollo puedan acceder fácilmente al mismo, se llegue a los destinatarios previstos y se aprovechen al máximo los beneficios de la evaluación." (p. 15).
	Respuesta a recomendaciones y seguimiento	"Las personas/entidades destinatarias de las recomendaciones deben responder de forma sistemática a las mismas y adoptar las medidas procedentes. Esto incluye una respuesta y un seguimiento formales por parte de la dirección. Las medidas acordadas han de ser supervisadas para garantizar que se rinde cuentas de su aplicación." (p. 15).
	Desarrollo de capacidad	"Deben maximizarse los efectos positivos del proceso de evaluación sobre la capacidad evaluativa de los socios para el desarrollo. Por ejemplo, una evaluación puede apoyar el desarrollo de capacidad gracias a la mejora de las destrezas y conocimientos en este ámbito, perfeccionando la gestión y estimulando la demanda y empleo de los resultados de la evaluación y fomentando un ambiente de aprendizaje y responsabilidad." (p. 7).
Reglas de pago	No existen estándares que puedan considerarse reglas de pago.	No existen estándares que puedan considerarse reglas de pago.

Tal como se puede observar en la tabla, al clasificar los estándares en los tipos de reglas del IAD, se ha procedido a hacer una subdivisión entre los estándares pertinentes -únicamente- para la situación de acción de diseño de la evaluación y los pertinentes para la situación de acción de ejecución de la evaluación. Esto es así, porque en estas dos etapas del proceso de evaluación se aplican distintas reglas, para distintos actores y se tienen distintos resultados potenciales.

Tal como se observa, en algunas ocasiones se repiten estándares dentro de las dos situaciones de acción, o entre los tipos de reglas en cada situación de acción. Los estándares de Participación de actores relevantes, Independencia de los evaluadores, Ética de la evaluación y Control de calidad se repiten en las dos situaciones de acción debido a que su naturaleza implica que son estándares que se deben tomar en cuenta a lo largo de todo el proceso de evaluación.

Dentro de la situación de acción de diseño de la evaluación, el estándar de Documentación relativa a la finalidad y las expectativas se clasifica como regla de información al tiempo que como regla de alcance. La razón de esto es que esta documentación en Colombia, denominada “Estudios Previos”, es el resultado potencial de los actores públicos en esta situación de acción; sin embargo, este documento debe ser comunicado al evaluador, sin lo cual, tal como se verá, no se podrá dar por terminada la etapa de diseño metodológico de la evaluación.

Por otro lado, en la situación de acción de ejecución de la evaluación se repite el estándar de Informe de la evaluación en reglas de información y en reglas de alcance. De manera similar al caso anterior, esto se explica porque el informe de evaluación es un resultado potencial de un actor (del evaluador), a la vez que es información que se debe comunicar y que deben tener los otros actores (particularmente la entidad sectorial evaluada) para poder llegar a cumplir el objetivo último de la evaluación (racionalización de las decisiones por medio del uso de información de la evaluación).

El estándar de respuesta a recomendaciones y seguimiento se repite en reglas de elección y de alcance, está dentro de la primera en el sentido en que es una acción necesaria el realizar seguimiento a la implementación de recomendaciones, y en la segunda razón, esta implementación es una dimensión fundamental del uso de la evaluación. Por su importancia se consideró procedente realizar esta subdivisión del estándar.

1.1.3.4. Papel de la interacción y coordinación de los actores en los procesos de evaluación

Considerando que la segunda variable independiente de la hipótesis central de la presente investigación es la coordinación entre el proceso de evaluación con el de la política pública evaluada, es importante aclarar conceptualmente esta variable, así como caracterizar su ubicación en la variable “patrones de interacción” de Ostrom. Derivado de la postura sistémica de esta autora, así como de su concepto de situación de acción, las interacciones o los patrones de interacción hacen referencia a las relaciones que se dan entre participantes dentro de la estructura de la situación de acción, por ejemplo, los intercambios de bienes y servicios, resolución de problemas, dominación de unos a otros, o luchas que se dan entre estos en un determinado espacio social o campo de acción (2015, pp. 51-53).

Esta segunda variable de la hipótesis también tiene como fuente de inspiración los postulados de Zittoun (2014) y Patton (2012), acerca de las particulares relaciones que deben darse entre los actores de política pública, específicamente entre evaluadores y los clientes o usuarios de la evaluación. Al ubicar el papel de la coordinación en las políticas públicas, Zittoun resalta que ninguna propuesta de política pública puede emerger sin el apoyo de cierta cantidad de stakeholders (2014, p. 137), esto es relevante en la medida en que el uso de una evaluación se

refleja de manera particular en (y puede medirse a partir de) la implementación de las recomendaciones de la evaluación.

Para entender el proceso en que una propuesta obtiene este apoyo, Zittoun aclara que el proceso definicional de estas propuestas (recomendaciones) no se puede desasociar de la construcción de acuerdos, los cuales son los que permiten ganar apoyo más allá de sus primeros proponentes. Los promotores de la propuesta, en este caso, los evaluadores que construyen las recomendaciones deben ser capaces de persuadir a los usuarios primarios de ser defensores activos de estas propuestas, para esto, deben contar con la habilidad de producir argumentos convincentes, y con la habilidad de redefinir la propuesta para -de este modo- hacerla apropiada a los nuevos actores (p. 138).

Para entender la interrelación entre construcción de acuerdos y definición de las propuestas de política, Zittoun utiliza el concepto de “ajuste mutuo” que Lindblom (1990) propone para comprender la coordinación que se da entre actores. Este concepto permite estudiar que un actor A busque persuadir a B, o que A y B busquen mutuamente persuadirse, sin considerar esto como simple manipulación, sino con atención en cómo el que intenta persuadir debe adaptarse a sí mismo considerando la situación en la que se encuentra, modificando sus palabras y actitudes en las discusiones (p. 134).

En esta perspectiva, la coordinación se caracteriza por la interdependencia de los actores (entre más interdependientes son los encargados de tomar decisiones, más fácil es la coordinación), que debe ser entendida como un proceso, el cual se da para alcanzar acuerdos con respecto a las decisiones (p. 134), y se refleja en múltiples ajustes que hacen posible que los intereses divergentes entre actores interdependientes se hagan compatibles (p. 135).

Estos ajustes mutuos pueden tomar dos formas distintas (p. 135): ajuste por medio de la adaptación, en donde un actor toma una decisión y el otro se adapta a esta; y ajuste por medio de la manipulación recíproca, en este caso su interacción puede tomar la forma de negociación, un trade-off, una discusión entre partes defensoras de distintas propuestas o una compensación, en esta forma existe un intercambio sin ser necesariamente una relación simétrica o balanceada. Este ajuste mutuo se desarrolla a partir de estrategias que los actores desarrollan para solucionar sus problemas (p. 136).

En consonancia y de modo más específico para temas de uso de la evaluación, Patton (2012) establece en su modelo Utilization-Focused Evaluation (U-FE) las características que debe tener el proceso de ajuste mutuo entre evaluadores y usuarios primarios de la evaluación. Estas características del proceso de coordinación como ajuste mutuo corresponden a la operacionalización de la segunda variable independiente de la hipótesis central de la presente investigación.

Desde la perspectiva de Patton, y en consonancia con la mayoría de los autores teóricos acerca de la evaluación (como se expuso anteriormente), las evaluaciones deben ser juzgadas a partir de su utilidad y uso actual (factor sin el cual no sería efectivas); por lo tanto, los evaluadores deben facilitar un proceso de evaluación y de diseño de cualquier evaluación, con una cuidadosa consideración de cómo todo lo que es hecho, desde el principio al fin, afectará el uso. El uso se refiere a como personas reales en el mundo real aplican los hallazgos de la evaluación y experimentan el proceso de evaluación (2012, p. 4). El informe de la evaluación y los hallazgos son simples outputs, el uso es el outcome deseado, lo que implica, entre otras cosas, trabajar con los usuarios primarios previstos de la información después de que los hallazgos y el reporte final han sido presentados (2012, p. 380) para -de este modo- actuar en los hallazgos.

El presupuesto psicológico que subyace al marco U-FE de Patton consiste en entender que es más probable que los usuarios previstos utilicen las evaluaciones, si ellos entienden y sienten apropiación del proceso de evaluación y sus hallazgos, y es más probable que ellos entiendan y sientan apropiación si están activamente involucrados, con lo cual el evaluador puede entrenar a los usuarios previstos en el uso, preparar el trabajo de campo para el uso, y reforzar la utilidad prevista de la evaluación en cada paso de su proceso (2012, p. 6). Esto también se propone de la siguiente manera:

> “Why bother? Why go to all trouble of trying to get these diverse people to agree? Why bother, indeed? The answer: increased credibility, user buy-in, and, ultimately, use.” (p. 46).

Para alcanzar estos acuerdos, Patton resalta la aplicación del estilo facilitador propuesto en el U-FE, el cual se aplica para todas las fases de interacción de los actores en el proceso de evaluación: parte de identificar a los usuarios previstos para involucrarlos en las decisiones de diseño, el establecimiento de prioridades, negociación de trade-offs, análisis y en la interpretación de los datos (p. 82). Estas interacciones entre evaluadores y usuarios previstos deben tener una naturaleza activa-reactiva-interactiva-adaptativa (p. 47), esto describe el modo en que se desenvuelve la toma de decisiones en el mundo real, y prescribe a los evaluadores a que de manera consiente y deliberada actúen, reaccionen, interactúen y se adapten en orden a incrementar su efectividad en su trabajo con los usuarios primarios. Este estilo requiere versatilidad, flexibilidad, creatividad, astucia política y sensibilidad.

El U-FE es activo en que de manera deliberada y calculada identifica a los usuarios previstos y se enfoca en preguntas útiles. Es reactivo en escuchar a los usuarios previstos y en responder a lo que ellos aprendan acerca de la particular situación en la que se desenvuelva la evaluación. Se es adaptativo en alterar las preguntas y el diseño de evaluación a la luz de la mayor comprensión de la situación y de las condiciones cambiantes (Patton, 2012, p. 47).

En términos normativos, y con respecto a la construcción colectiva de acuerdos mediante discusiones en las cuales se debería ver reflejada la coordinación como ajuste mutuo, pueden sintetizarse del modelo U-FE los siguientes lineamientos:

1. Hallar e involucrar a los usuarios primarios previstos en todo el proceso de evaluación (2012, pp. 61-68). Usuarios deben tener: interés, conocimiento, sinceridad, conexión con interesados con influencia/poder de implementación, credibilidad, abiertos al aprendizaje, comprometidos y disponibles para la interacción a lo largo del proceso de evaluación (2013, p. 4).
2. Evaluadores y usuarios primarios deben formar un equipo de trabajo para la toma de decisiones para todo el proceso de evaluación (2013, p. 4).
3. Escuchar atentamente (y reaccionar/ser reactivo) a las preocupaciones prioritarias de los usuarios para ayudarlos a identificar preguntas importantes. Estar en sintonía con las preocupaciones de los usuarios y no permitir que el interés del evaluador domine o controle el proceso de ajuste de prioridades (p. 9).
4. El evaluador trabaja con usuarios principales previstos para obtener claridad y acuerdo sobre los objetivos prioritarios destinados a lograr el objetivo de la evaluación: usos previstos por los usuarios previstos (2012, p. 13). Se busca establecer prioridades y resolver conflictos sobre propósitos en competencia, evitando ambigüedad o confusión acerca de las prioridades (2013, p. 7).
5. Mostrar y revisar con los usuarios previstos el menú de opciones de preguntas que definen los enfoques de evaluación (p. 9).
6. Discutir con los usuarios previstos, de manera útil y práctica, los debates metodológicos que sean relevantes y que afecten las decisiones metodológicas (p. 12). Negociar trade-offs entre los ideales metodológicos y lo que realmente se pude implementar debido a las inevitables limitaciones de tiempo y recursos, negociar criterios de calidad metodológica y sobre lo que constituye evidencia creíble (p. 11).
7. Adaptar los métodos en respuesta a las condiciones cambiantes a medida que se desarrolla el proceso de evaluación (p. 11).
8. Involucrar a los usuarios primarios en la interpretación de los hallazgos: Trabajar con los usuarios en pensar lo que es significativo y considerar explicaciones alternativas para los hallazgos antes de construir conclusiones; Involucrar a los usuarios en los juicios evaluativos (p. 14).
9. Involucrar a los usuarios previstos en la construcción de las recomendaciones: discutir el costo, los beneficios y los desafíos de implementar recomendaciones; enfocarse en acciones bajo el control de los usuarios previstos y de aquellos en los que podrían influir (p. 15).
10. Desarrollar un plan de seguimiento a la implementación de recomendaciones en conjunto con los usuarios primarios (p. 16).
11. Buscar proactivamente la utilización de la evaluación: vigilar y reaccionar ante el surgimiento de oportunidades que refuercen la relevancia de los hallazgos (p. 16).

Como se observa, todos estos lineamientos constituyen un modo particular de coordinación a partir del ajuste mutuo los actores, por lo general toman la forma de ajuste a partir de manipulación recíproca, con excepción de la determinación conjunta del uso previsto, el cual, al depender de la demanda de información de los usuarios, los evaluadores deberán adaptarse a esto. Esta coordinación, como aspecto central del U-FE, implica una comunicación continúa[31] (en consonancia con la coordinación como proceso de Lindblom), que es desarrollada a partir del establecimiento de un grupo de trabajo para la toma de decisiones en el proceso de evaluación.

Este proceso implica trabajar con los usuarios primarios previstos después de que los hallazgos y el reporte final han sido presentados (2012, p. 380), particularmente a través del seguimiento a la implementación de las recomendaciones. Para esto es necesario que los planes de trabajo de las evaluaciones (incluyendo los contratos) no terminen con la producción del reporte final, sino que se destine tiempo y dinero (2012, p. 381) al seguimiento de la implementación de las recomendaciones.

1.1.3.5. Operacionalización de la efectividad como uso instrumental de la evaluación

En esta investigación se aborda la efectividad de los procesos de evaluación a partir del uso de la evaluación, se ha manifestado que esto se refiere a cómo personas reales en el mundo real aplican los hallazgos de la evaluación y experimentan el proceso de evaluación (Patton, 2012), pero esta definición sigue siendo demasiado abstracta. Diversos autores recientes han establecido que el uso de la información es una variable manejable que permite estimar los impactos de reformas administrativas en el comportamiento humano (Moynihan, Pandey y Wright, 2012, p. 2; Van de Walle y Van Dooren, 2008). En este tema existen múltiples tipologías. Patton (2012, p. 115; p. 140), por ejemplo, establece la existencia de seis tipos de usos distintos que se le puede dar a las evaluaciones, estos usos son: 1. Sumativo; 2. Formativo/orientado al mejoramiento; 3. Rendición de cuentas/control; 4. Monitoreo; 5. Generación de conocimiento; 6. Desarrollo (adaptación a condiciones complejas, emergentes y dinámicas); y 7. Usos de proceso, en los cuales se busca que la organización evaluada aprenda a producir conocimiento evaluativo.

Moynihan, Pandey y Wright identifican como posibles usos de la información el uso intencional, donde se utiliza para modificaciones administrativas del programa, o para ubicación de recursos según requiera el mejoramiento de metas; el uso pasivo,

[31] En cuanto a este punto, Hofstetter y Alkin (2003), citados por Patton, proponen:“The evaluator could enhance use by engaging and involving intended users early in the evaluation, ensuring strong communications between the producers and users of evaluations, reporting evaluation findings effectively so users can understand and use them for their purposes, and maintaining credibility with the potential users” (2003, p. 216).

en donde los actores hacen lo mínimo requerido sin realmente usar los datos resultantes del sistema; el uso político en donde se usan los datos para argumentan a favor del valor y la legitimidad de un programa y así obtener recursos; y el uso perverso, el cual es parecido al político con la adición de que se manipulan los datos de manera contraria al rendimiento real a partir de mecanismos como la manipulación de metas (2012, p. 3).

Sandison (2006, pp. 7-8) incluye dentro de su tipología de usos de la información el uso instrumental, correspondiente a la directa implementación de los hallazgos y las recomendaciones; uso conceptual, en donde la información de las evaluación es introducida en las organizaciones en la forma de nuevas ideas y conceptos deliberados, y se van desarrollando con el tiempo; usos de proceso (en consonancia con Patton), en donde se da un aprendizaje y fortalecimiento de capacidades en la aplicación de herramientas de monitoreo y evaluación; y uso de legitimación, en el cual la evaluación se orienta a confirmar, sustentar o corroborar una decisión o comprensión que la organización ya tiene. La autora complementa lo anterior, con una tipología de no usos (p. 9), en la cual incluye: uso ritual, en donde la evaluación sirve únicamente a propósitos simbólicos de requerimientos legales o para representar un adecuado grado de rendimiento en ejercicios de rendición de cuentas; el Mal-uso, en donde se suprimen, subvierten o se distorsionan los hallazgos para servir a una agenda sesgada; y el no-uso, en donde la evaluación es ignorada porque los usuarios le encuentran poco o ningún valor a los hallazgos

Otro autor que realiza una tipología al respecto es Balthasar (2009, pp. 10-11), en esta incluye: uso instrumental, el cual puede ser documentado y se refiere al uso propuesto y directo de los hallazgos por parte de los tomadores de decisión; uso conceptual; el uso conceptual, con una conceptualización idéntica a la de Sandison; uso simbólico en donde los tomadores de decisión usan los hallazgos para confirmar su perspectiva o para legitimarse a sí mismo, muy similar al uso político de Sandison; uso de proceso (en consonancia con Patton); y uso general, en donde se valoraría el beneficio general obtenido a partir de la evaluación, lo que incluiría una valoración de los tipos particulares de utilización de la evaluación. Una tipología muy similar también es utilizada por Appleton-Dyer et al. (2012), la cual incluye los primeros cuatro tipos de uso descritos por Balthasar.

Estas tipologías tienen muchos elementos en común, a pesar de que a algunos tipos de usos se les da nombres distintos, o que se agregan algunos tipos adicionales, en general se especifican cuatro usos principales. El uso instrumental guarda muchas similitudes con el uso intencional, o a los usos formativos y sumativos descritos por Patton. Esta perspectiva de uso en el sentido de implementación ha sido reconocida como el aspecto más difícil pero crucial del uso de la información para llevar a cabo reformas en las organizaciones, en el sentido en que requiere cambios en el comportamiento y la ampliación y profundización del alcance de la información a tener en cuenta (Saliterer y Korac, 2013, De Lancer and Holzer, 2001; Hatry, 2002; OECD, 2005; Torres et al., 2011; Van Dooren, 2008). Adicionalmente,

en el sentido en que lleva a cambios en la política, recursos, sistemas o práctica operacional, permite observar cambios tangibles (Sandison, 2006, p. 5; Kroll, 2014, p. 5; Peck y Gorzalski, 2009).

En esta investigación se va a restringir el concepto de uso de la evaluación a su uso instrumental. Esto es así por razones prácticas y teóricas. Por una parte, tanto el componente de evaluaciones estratégicas de SINERGIA (tal como se ve en el análisis de reglas de alcance del capítulo 2), como los estándares profesionales de la OCDE (OCDE, 2010, p. 7) tienen objetivos claramente orientados y restringidos a un uso instrumental de la información. Por lo que, utilizar este tipo particular de conceptualización ayudaría a poner a prueba la Política de Evaluación de Política, y la principal institucionalización a nivel internacional en occidente acerca de la calidad de los sistemas y procesos de evaluación.

Por otra parte, en términos teóricos el uso instrumental puede ser traducido a términos de Ostrom. En el IAD las variables en las que se reflejaría un uso instrumental de la información son: los participantes de la situación de acción de la evaluación (particularmente, la dimensión denominada "modelo mental"[32], la cual corresponde -en este caso- a la representación de la política pública y de las instituciones que determinan su transformación, que tienen los usuarios previstos); y la institucionalidad de la política evaluada, en la cual se da un cambio con base en la selección de las acciones de los participantes producto de la modificación de su modelo mental.

Esto ayuda a operacionalizar la variable dependiente de la hipótesis del presente estudio. El cambio del modelo mental de los participantes de las situaciones de acción de las políticas evaluadas es posible medirlo observando el grado de la implementación de las recomendaciones de la evaluación, con la condición de que estas recomendaciones correspondan a nuevas opciones de acciones de los participantes de las situaciones de acción de las políticas evaluadas. En contraposición, si se puede observar que estas recomendaciones son acciones que ya se consideraban como necesarias por parte de los funcionarios de las entidades evaluadas antes de tener disponibles los hallazgos de la evaluación, no habría cambio en el modelo mental, la recomendación correspondería a la justificación de una decisión previa (el denominado uso simbólico o político), no se mejoraría el proceso de toma de decisiones, y la implementación de esta acción no reflejaría efectividad del proceso de evaluación. Estas recomendaciones van a ser descartadas como casos de efectividad, porque este "uso simbólico" no permite utilizar la información recolectada, de tal manera, que permita abordar los objetivos

[32] Es importante resaltar que, acompañando a la implementación de recomendaciones, el uso también debe tener como componente lo que Ostrom denomina una "revisión de los modelos mentales" de los participantes, la cual, solo es posible a partir de la consideración de información sobre los resultados de acciones previas y sobre los posibles resultados de nuevas acciones disponibles –recomendaciones- (ver esquema del holón "participante"), de esta consideración se desprenderán las subsecuentes acciones a elegir por los participantes de las situaciones de acción de la política evaluada (que también pueden ser identificados como usuarios previstos en términos de Patton).

de la presente investigación, y no está en consonancia con las teorizaciones que se quieren poner a prueba.

Finalmente, para observar efectividad, es fundamental evidenciar, en términos de Ostrom, que los funcionarios de la política evaluada seleccionan acciones en concordancia con el cambio de su modelo mental. Esto se puede ver reflejado a partir de cambios en la institucionalidad de la política pública evaluada que vayan en dicha dirección. Es necesario aclarar que a pesar de que las recomendaciones producto de las evaluaciones son outputs de arenas de acción de nivel operativo de la política de evaluación de políticas, al mismo tiempo son inputs de las arenas de acción de los niveles de elección colectiva y constitucional de la política evaluada[33], en donde las interacciones de los participantes siempre producen resultados que consisten en cambios institucionales.

En síntesis, todo esto significa que para medir el uso de la evaluación, además de ver el cambio del modelo mental de los usuarios, es necesario observar el rastro que queda de una selección de acciones en concordancia con ese cambio de modelo, al interior de la institucionalidad de la política pública. Cambios institucionales orientados hacia la implementación de las recomendaciones del proceso de evaluación, son la mejor evidencia de dicho rastro.

1.2. Síntesis del proceso analítico

Ostrom recomienda utilizar estudios minuciosos de caso para examinar las relaciones complejas entre acciones y resultados en la estructura de la situación de acción, los cuales, una vez se han construido con todo el conjunto de variables propuestas, permiten proceder hacer un estudio comparado de casos (2015, p. 76). Con esta comparación de casos se pueden determinar las relaciones causales de interés, abordar el objetivo principal de este estudio, e incluso se pueden recomendar fortalecimientos institucionales en aspectos específicos.

El muestreo realizado, en consonancia con el método de análisis comparado, son 8 casos de procesos de evaluación de política pública realizados por la Dirección de Seguimiento y Evaluación de Políticas Públicas (DSEPP) del Departamento Nacional de Planeación realizados entre el 2014 y el 2017. Las razones de seleccionar estos casos son las siguientes: por un lado, se requería que fueran procesos de evaluación con al menos dos años transcurridos desde su terminación para así dar tiempo suficiente para observar la implementación o no de las recomendaciones. Por otro lado, no podían ser procesos de evaluación muy

[33] Las recomendaciones nunca influyen directamente a las arenas de nivel operativo de la política pública evaluada, lo cual se observa en que las recomendaciones nunca quedan como reglas, normas o estrategias en sí mismas, sino que deben ser procesadas en arenas de acción de la política pública evaluada.

antiguos debido a la alta rotación de personal que caracteriza la burocracia en Colombia, situación que fue evidenciada con el fracaso en la búsqueda de contacto con los usuarios previstos (funcionarios de entidades sectoriales) que participaron en diversos procesos de evaluación, ante lo cual hubo que ir descartando varios casos.

Para determinar en qué medida la política pública de evaluación de políticas en Colombia permite confirmar o falsear la incidencia de la teorización seleccionada en el uso de la información como modo de racionalización en la toma de decisiones de política pública, se consideran necesarias las siguientes tareas:

1. Describir a profundidad el grado de fallas de implementación que tienen los diversos casos de procesos de evaluación de política pública a comparar.
2. Describir a profundidad el grado de efectividad que tienen los diversos casos de procesos de evaluación de política pública a comparar.
3. Brindar una estimación acerca del grado de efectividad actual que tiene la política pública de evaluación de políticas públicas.
4. Analizar la importancia relativa de los diversos estándares de calidad para la generación de efectividad por parte de los procesos de evaluación de política pública.
5. Analizar la importancia relativa que tienen los componentes de la coordinación, entre el proceso de evaluación con el de la política pública evaluada, para la generación de efectividad por parte de los procesos de evaluación de política pública.

Las primeras dos tareas remiten a un ejercicio de medición de cada caso que será presentado en los distintos capítulos de casos. La tercera tarea remite a un análisis descriptivo conjunto de la tendencia que presentan los casos con respecto a la efectividad de los procesos evaluativos. Las últimas dos tareas son más complejas y se pretenden realizar a partir del contraste de resultados entre dos métodos de análisis: El método comparado de similitud, y el método comparado de diferencia.

El análisis de similitud o también llamado análisis de condiciones necesarias se realizará a partir de dos pasos (Pérez, 2007, p. 20): primero se seleccionarán todos los casos positivos con respecto a la variable dependiente, es decir, en todos los que existe efectividad (uso de la información como modo de racionalización en la toma de decisiones de política pública), y luego, se procederá a identificar las condiciones causales que son comunes a todos ellos, esto a partir de la regla de interpretación en donde si una condición X es necesaria para que exista Y, entonces todos los casos en los que exista Y deben ofrecer también la presencia de X.

Esto permitirá identificar aquellos lineamientos provenientes de las teorizaciones de evaluación que son necesarios para que exista la variable dependiente, así como los que no son necesarios, permitiendo así, comprobar o falsear las teorías sobre evaluación. Sin embargo, solo tener casos positivos en donde exista siempre la variable dependiente y, por ende, sin varianza en esta variable, no permite hacer

inferencias causales válidas por sesgo de selección (p. 9). En este análisis, eso significa que puede haber casos en donde no existe la variable dependiente, pero en los que si se dan los factores causales determinados como necesarios. Casos que no fueron tomados en consideración y, por ende, no se podría contra argumentar frente a su existencia.

Para solucionar esto, los resultados del análisis comparado de similitud se contrastarán con el análisis comparado de diferencia del cual también se obtendrá evidencia para las conclusiones acerca de la posibilidad de falsear o comprobar las teorías. En este análisis se procederá a seleccionar tanto casos positivos como negativos en el estado de la variable dependiente, luego los factores causales se estudiarán bajo la siguiente lógica: "… todos los casos que presentan el resultado de interés deben también compartir una característica explicativa común (con suerte, aquella característica anticipada por nuestra hipótesis), mientras que los casos negativos deben carecer de este atributo" (p. 10). Como se observa, se valora como factores causales aquellas variables que coinciden en su varianza con la variable dependiente.

A efecto de cumplir el objetivo central, se resaltarán como comprobadas aquellas variables que coincidan como efecto causal tanto del análisis comparado de similitud como con el de diferencia, y se falsearán, o por lo menos podrán ser consideradas como irrelevantes, para el caso colombiano, aquellas variables que resultaron innecesarias en el primero e incongruentes con la varianza de la variable dependiente en el segundo.

Por último, esta metodología y sus resultados serán complementados con un estudio de caso general de toda la política pública de evaluación de políticas en Colombia. En este estudio de caso, se procederá a verificar la correspondencia entre los estándares de calidad clasificados en tipos de reglas, con el actual diseño institucional de la política pública en cuestión (1ra variable independiente), luego se caracterizará la existencia de los componentes de la coordinación como ajuste mutuo delimitados por Patton (2da variable independiente), y se relacionarán estos resultados con una primera medición de la efectividad de la política pública en la cual se tomarán como fuente de información entrevistas a los funcionarios de la DSEPP-DNP y a evaluadores contratados por esta. El propósito de este caso general es observar la pertinencia y posibilidad de estudiar las variables seleccionadas en este modelo, en el caso colombiano, y brindar una primera impresión acerca de la efectividad de la política pública.

La contrastación de estos resultados, junto con la aplicación de los dos métodos comparados ayudarán al surgimiento de resultados más rigurosos. Este estudio de caso general se presentará en el siguiente capítulo, y ayudará a enmarcar el estudio al caracterizar las instituciones dentro de las cuales se desarrollan los contratos de los procesos de evaluación seleccionados como muestra de la metodología comparada.

1.3. Conclusiones

Como balance final de este capítulo, se debe resaltar que los conceptos del IAD de Ostrom permiten abordar de manera simplificada todo el complejo conjunto de variables que implica poner a prueba tanto los estándares de calidad de la OCDE, como la coordinación del U-FE de Patton. Si bien este último también fue enmarcado conceptualmente a partir del enfoque pragmático de Zittoun y del concepto de coordinación como ajuste mutuo de Charles Lindblom.

Todos los conceptos de estos autores comparten supuestos importantes sobre los cuales se puede realizar una armonización, particularmente pueden ser ubicados dentro de la corriente postpositivista del estudio de las políticas públicas, por lo que comparten supuestos ontológicos, epistemológicos y metodológicos. Con respecto a este último, en palabras de Guba (1990), el postpositivismo aboga por una postura metodológica que resalta el multiplismo crítico (pensado como una forma de triangulación), en donde se entiende que los hallazgos de una investigación deben sustentarse a partir de tantas fuentes (datos, investigadores, teorías y métodos) como sea posible, confiando en que esta heterogeneidad de fuentes haga menos probable la distorsión por las interpretaciones que se realicen (p. 21). Esta postura pretende ser satisfecha a partir de combinar estudios de caso a profundidad, análisis comparado de similitud y análisis comparado de diferencia.

2. Estudio de caso general de la Política Pública de Evaluación de Políticas en Colombia

2.1 Análisis institucional

El análisis de este capítulo está orientado a servir como punto de verificación de la pertinencia de estudiar cada uno de los estándares profesionales de calidad identificados en el modelo analítico, para aclarar el marco institucional en el que se desarrollan los casos específicos a comparar y, por último, para brindar una primera apreciación del grado de efectividad de la Política de Evaluación de Políticas en Colombia. En términos de la verificación antes mencionada, es necesario tener claridad que la existencia de los estándares de calidad, y lineamientos coherentes con el marco U-FE dentro de la normatividad legal o técnica de la Política Pública de Evaluación de Políticas en Colombia, no es evidencia de una real implementación de estos elementos.

En este capítulo se profundiza en la normatividad jurídica y técnica de la política en cuestión, particularmente en toda aquella institucional que es transversal a todos los procesos de evaluación/contratación de evaluadores de la DSEPP (por lo que no se toma en cuenta institucionalidad particular de los casos seleccionados, como la que se puede encontrar en pliegos de condiciones o contratos). La estructura de este análisis institucional está dada por los tipos de reglas descritos por Ostrom, y que determinan la estructura de las situaciones de acción en general. Se omite las reglas de pago debido a que no existe ningún estándar orientado en el sentido de dicho tipo de reglas.

2.1.1 Reglas de alcance

Tal como se manifestó anteriormente, las reglas de alcance corresponden a la delimitación de los resultados potenciales que pueden ser producidos, y a las

acciones vinculadas a esos resultados específicos. Este tipo de reglas pueden encontrarse en varios lugares. En primer lugar, del Decreto 1290 de 2014 “Por el cual se reglamenta el Sistema Nacional de Evaluación de Gestión y Resultados - SINERGIA” se manifiesta que la evaluación de gestión y resultados es uno de los “más eficaces instrumentos” para la modernización del sector público siendo este el resultado potencial más general (o abstracto) de la política. En concordancia con la perspectiva del U-FE, así como del consenso académico acerca de la utilidad de la evaluación de políticas, tanto en el Decreto 1290 de 2014, art. 1), como en el 1082 de 2015 (art. 2.2.7.1.1.1.) se aclara que el objetivo de SINERGIA es “generar información de calidad para la toma de decisiones que permitan mejorar la efectividad de la formulación y ejecución de políticas”.

Particularmente hablando del componente de evaluación[34], las reglas de alcance encontradas giran alrededor de generar información que sirva como insumo para el mejoramiento de las intervenciones públicas (Decreto 1290 de 2014, art. 21); generar evidencia que sirva como insumo para mejorar el diseño, la implementación, la ejecución y los efectos de las intervenciones públicas (art. 22; DSEPP-DNP, 2014a, p. 34); propender por el uso de los resultados y las recomendaciones de las evaluaciones como insumo para la toma de decisiones (Decreto 1082 de 2015, Art. 2.2.7.3.1.4.).

Las evaluaciones legalmente están estructuradas como proceso (art. 2.2.7.3.2.2), en el cual se parte de una fase de selección, se continua con el diseño, sigue su ejecución y se finaliza con la socialización y uso de la evaluación. En esta última fase, se establece que se deben realizar las gestiones necesarias para que las entidades ejecutoras de las intervenciones evaluadas utilicen los resultados e incorporen las recomendaciones en sus procesos de toma de decisiones (art. 2.2.7.3.2.12.).

La Dirección de Seguimiento y Evaluación de Políticas Públicas (DSEPP) del DNP es la encargada de liderar la implementación de la política de evaluación. En sus guías metodológicas también caracterizan el uso/efectividad de la evaluación como el que los tomadores de decisión “… de manera informada, formulen acciones de mejora para la intervención evaluada” (2014a, p. 34); servir como insumo para la retroalimentación del ciclo de la política pública (DSEPP-DNP, 2012a, p. 53). También se aclaran particulares usos dependiendo del tipo de evaluación que se realice:

[34] Puesto que Sinergia también cuenta con otro componente adicional a las evaluaciones focalizadas, el cual se centra en el seguimiento a metas de gobierno.

Tipo de evaluación	Regla de alcance: uso
Procesos	- "Identificar y modelar los procesos y actividades de un programa. - Determinar si la cadena de producción del programa, compuesta por los procesos, se lleva a cabo según su diseño. - Generar esquemas de gestión de calidad que localizan los problemas, defectos y errores en los procesos y actividades del programa. - Crear herramientas de control y ajuste que garanticen que los procesos y actividades sean eficientes y eficaces." (2012 a, p. 35).
Productos	Identificar "… mejoras para incorporar al programa en temas de gestión de insumos, actividades o productos para incrementar la eficiencia y eficacia en la entrega de estos. Posteriormente, se debe hacer un plan de implementación de estas mejoras con las respectivas responsabilidades y cronograma para llevarlas a cabo" (p. 43).
Ejecutiva	- "Tener una visión general del funcionamiento del programa en un corto plazo. - Establecer la calidad del diseño del programa. - Identificar si el enfoque del programa es correcto. - Comprobar la calidad de los procesos. - Determinar la calidad en la entrega de los productos y descubrir cuellos de botella. - Constatar la cobertura del programa" (p. 59).
Resultados	- "Documentar el funcionamiento del programa, especialmente la forma en la que los productos afectan a los resultados. - Mejorar los programas. - Apoyar la planeación anual o de largo plazo. - Guiar presupuestos y justificar la asignación de recursos. - Sugerir metas de resultados. - Enfocar la atención de hacedores de política en puntos críticos de los programas" (p. 69),
Impacto	- "Atribuir impactos a un programa y examinar si está generando los efectos deseados. - Tener evidencia concluyente para la toma de decisiones de política, así como en la priorización y justificación en el uso de los recursos. - Estudiar aspectos de un programa para cambiar su diseño y mejorar su efectividad.

	- Comparar el impacto de un mismo programa para diferentes tipos de población o diferentes contextos. - Identificar efectos no previstos en los beneficiarios, ya sean positivos o negativos" (p. 79).
Costo-beneficio	- "Estimar y cuantificar los costos de un programa. - Estimar y cuantificar los beneficios de un programa. - Proporcionar una medida de rentabilidad de un programa. - Elegir la mejor alternativa entre programas similares" (p. 106).
Institucional	- "Estudiar tanto los aspectos organizacionales como los estructurales que intervienen en el logro de los objetivos del programa. - Identificar la capacidad institucional actual para lograr ciertos objetivos o entregar determinados productos. - Encontrar puntos débiles e implementar cambios en la organización. - Formular actividades, procedimientos o nuevas estructuras institucionales para optimizar la entrega de productos. - Comparar las diferentes organizaciones que prestan el servicio. - Proponer correctivos en las relaciones organizacionales. - Mejorar la entrega de bienes y servicios." (p. 128).

Como se puede observar, todos los tipos de evaluación a los que la DSEPP les ha determinado sus usos, tiene uno o varios usos centrados en la implementación de decisiones producto de la información de la evaluación (con excepción de la evaluación ejecutiva, la cual no es seleccionada en el muestreo de la presente investigación). Esto está en consonancia con la operacionalización de la efectividad de los procesos de evaluación de la presente investigación.

En síntesis, con relación a las reglas de alcance y su concordancia con los estándares profesionales de calidad se puede observar que, en efecto, se establece el estándar de oportunidad, pertinencia y utilización. Esto es así, porque se evidencia que el marco institucional de las evaluaciones se orienta a que estas satisfagan las necesidades de los usuarios previstos, a que construyan sus conclusiones y recomendaciones de forma clara, pertinente, focalizada y orientada a su puesta en práctica, para así lograr los objetivos de aprendizaje.

Por su parte, frente al estándar de respuesta sistemática a las recomendaciones y seguimiento de estas, se evidencian lineamientos dentro de la normatividad técnica de la DSEPP. Dentro de la Guía metodológica para el Seguimiento y la Evaluación a Políticas Públicas (DSEPP-DNP, 2014a, p. 48) se describe como tercera actividad de la fase de socialización y uso de los resultados de la evaluación, la elaboración de la Matriz de Gestión de Recomendaciones.

Esta matriz permite la revisión de los resultados y recomendaciones diseñadas por los consultores contratados por la DSEPP para la elaboración de las evaluaciones. La implementación de estas recomendaciones: "… permitirá fortalecer y replicar las acciones exitosas de la intervención, resolver los cuellos de botella detectados y resaltar otros elementos relevantes sobre los cuales se debe desarrollar alguna acción" (p. 48). Su elaboración está a cargo del Comité Técnico de Seguimiento a la Evaluación[35], y tiene los siguientes lineamientos para su elaboración:

- "Definir la ruta de uso de la evaluación, es decir las acciones que se realizarán en torno a las propuestas realizadas por la evaluación y difusión con actores adicionales, entre otros." (p. 48).
- "Establecer el orden de prioridades para enfocar los esfuerzos de implementación en las recomendaciones más relevantes para la entidad ejecutora" (p. 48).
- "Diligenciar la propuesta inicial de la MGR, a la luz de esta definición de recomendaciones prioritarias o de su totalidad, para discutir con la entidad ejecutora de la intervención" (p. 48).

Los nombres de las columnas de la matriz que sirven para caracterizar y gestionar a las recomendaciones son:

- Orden o factor de prioridad.
- Recomendación.
- ¿Acoge la recomendación? Si/No.
- Observaciones.
- Acciones por desarrollar.
- Actor responsable.
- Tiempo de implementación

El Comité de Seguimiento debe revisar la matriz de recomendaciones, con lo cual se verifica el alcance de los compromisos adquiridos, y todo se diligencia con una directriz de buscar consensos (p. 49). Luego de que se aprueba esta matriz, se debe elaborar un acta cuyos elementos son (p. 49):

- Descripción de acciones realizadas para difundir los resultados y recomendaciones de la evaluación.
- Proceso de elaboración de la matriz, el número de recomendaciones acogidas y rechazadas.
- La descripción de la ruta de uso de la evaluación, la cual define acciones en las que la evaluación podría ser tenida en cuenta como insumo.

[35] Grupo de supervisión al desarrollo de la evaluación conformado por funcionarios del DNP y de las entidades sectoriales evaluadas. Su conformación, rol y actividades se detallan más adelante en la presentación de las reglas de frontera, de posición y de elección.

Se evidencia que la implementación de recomendación, por lo menos normativamente, toma un carácter muy importante dentro de la definición del uso de la evaluación. En los lineamientos de esta Matriz de Gestión de Recomendaciones incluso se aclara que para las recomendaciones que se definan como “no acogidas”, el actor responsable debe justificar en la casilla de observaciones las razones por las cuales es rechazada (p. 48).

A pesar de que hasta el momento se puede considerar que el estándar relacionado con las recomendaciones está incluido en la normatividad técnica de la DSEPP, así como del alcance limitado de verificación normativa que se le ha dado al presente capítulo, en este punto es necesario hacer una aclaración debido a su importancia. Tanto las matrices de gestión de recomendaciones como las actas de su aprobación fueron solicitadas al DNP, frente a las primeras se respondió (Respuesta a derecho de petición N° 20183600708761) únicamente con una de las matrices solicitadas, frente a los otros 9 casos de evaluación se anexó un nuevo tipo de formato denominado “matriz de uso” con el cuál se reemplazó la Matriz de Gestión de Recomendaciones[36] desde el 2015 (un año después de la elaboración de la guía en la que se establecen y publican los lineamiento de esta última).

Por su parte, frente a las actas se contestó: “… no se establecen acta para la aprobación de la Matriz de Gestión de Recomendaciones, por lo tanto, la Entidad no cuenta con estos insumos”. Por lo que no existe acta para ningún caso (incluyendo en el que si se hizo Matriz de Gestión de Recomendaciones).

Frente al nuevo tipo de formato de matriz que se está utilizando, en este se cuenta con los siguientes nombres de columnas de caracterización y gestión de recomendaciones:

- Categoría de uso
- Subcategoría de uso
- Recomendación
- Acciones DNP
- Nombre de entidad responsable
- Comentarios adicionales

El estándar de respuesta sistemática a las recomendaciones y seguimiento a las mismas establece que se deben adoptar medidas procedentes a las recomendaciones, lo que incluye una respuesta, un seguimiento y una supervisión por parte de la dirección de la evaluación que garantice que se rinde cuentas de la

[36] Puntualmente se respondió: “Adjunto a la respuesta se remite en formato comprimido la matriz de gestión de recomendaciones de la evaluación de operaciones de la Política de atención humanitaria a población en situación de desplazamiento. Con respecto a las otras evaluaciones relacionadas en su solicitud, le informamos que este documento fue reemplazado por las matrices de uso, por consiguiente, anexo le remitimos la información correspondiente.”

aplicación de estas. Los componentes de Matriz de Gestión de Recomendaciones brindaban lineamientos puntuales acerca de una respuesta formal, así como de la determinación de los elementos puntuales a los cuales se les debía hacer seguimiento y supervisión.

Por su parte, la nueva matriz de uso únicamente considera acciones del mismo DNP, y una columna de comentarios adicionales (que siempre se llenó con N/A en todos los campos de las evaluaciones seleccionadas y que se eliminó en 2017 de este formato) por lo que tiene un carácter unilateral en su utilidad, esto a pesar de la fuerte conexión que se establece entre uso de la evaluación y la implementación de las recomendaciones por parte de las entidades sectoriales en la normatividad técnica.

En este sentido, ninguno de los componentes del estándar de respuesta a las recomendaciones es satisfecho con esta nueva herramienta de uso que remplazó (en palabras del DNP), la herramienta puntual para la respuesta y el seguimiento a las recomendaciones producidas por las evaluaciones. El único estándar al cual no se le encontró normatividad jurídica o técnica relacionada perteneciente a la categoría de reglas de alcance, corresponde al estándar de desarrollo de capacidad. Sin embargo, no se descarta que este pueda estar plasmado dentro de la institucionalidad específica en los contratos de cada proceso de evaluación.

2.1.2 Reglas de posición y de frontera

Las reglas de posición determinan el rol o los tipos de participantes que tendrá la situación de acción, mientras que las de frontera establecen el número de participantes, sus atributos, sus recursos, si pueden entran libremente, y sus condiciones para salir de una determinada situación de acción. Las reglas de posición crean posiciones, mientras que las de frontera “influyen en la forma en la que se asignan o quitan posiciones” (Ostrom, 2015, p. 258) a los participantes de la situación. Se presentan juntas porque se evidencia que todas las reglas de frontera son definidas al tiempo que varias de las reglas de posición[37].

En el Decreto 1082 de 2015 se establece que los actores del proceso de evaluaciones estratégicas son los ministerios, departamentos administrativos y sus entidades vinculadas y adscritas (art. 2.2.7.3.1.3.). Esto no se refiere a todo el conjunto de participantes de las situaciones de acción que componen el proceso de evaluaciones, sino únicamente a los beneficiarios de la política.

[37] Esto es algo que Ostrom manifiesta que puede suceder (2015, p. 262).

Por su parte, al DNP se le da el rol de planear, diseñar y organizar los sistemas de evaluación de gestión y resultados en la Ley 152 de 1994 orgánica del Plan de Desarrollo (art. 29), por lo cual, tiene dentro de sus competencias el señalar los responsables, los términos y las condiciones para realizar las evaluaciones (Decreto 1290 de 2014).

Las evaluaciones son lideradas por dos comités: el Comité de Diseño de la Evaluación y el Comité Técnico de Seguimiento de la Evaluación. El Comité de Diseño se estable en una etapa previa al de Seguimiento, y hay un esfuerzo por parte de la DSEPP de que los mismos actores que participan en el primero sigan presentes en el segundo (como se verá más adelante).

Tanto el Comité de Diseño como el Comité de Seguimiento deben ser conformados por el DNP, con delegados de la DSEPP, de las direcciones técnicas del DNP que tengan competencias en los temas abordados por la evaluación, y por delegados de la entidad ejecutora de la intervención a evaluar (Decreto 1290 de 2014, art. 29 y art. 32). Estos comités son los escenarios de coordinación en los cuales se tomarán las más importantes decisiones con respecto a las evaluaciones.

Adicional a estos comités, a cada proceso de evaluación se le asigna un coordinador, el cual, es un funcionario o contratista de la DSEPP que debe cumplir con requisitos (DSEPP-DNP, 2014a, p. 36) como: conocimiento sectorial; conocimiento en tipos de evaluación; capacidad de liderazgo y coordinación; capacidad de trabajo en equipo; capacidad analítica y de interpretación de resultados para generar recomendaciones; conocimiento en metodologías de recolección y análisis de información.

El Comité de Diseño es el que identifica el recurso humano necesario para la evaluación (p. 41), en donde se determinan los perfiles apropiados y el número de personas que debe tener el equipo de evaluación (el contratista/evaluador), esto incluye los niveles académicos y la experiencia laboral de cada una de las personas que se consideró necesarias, y este perfil queda plasmado como regla de frontera (requisito para presentarse como candidato a realizar la evaluación) para las firmas consultoras o los equipos de profesionales que se presenten.

En la situación de diseño de la evaluación hay tres roles o posiciones: el coordinador de la evaluación de la DSEPP-DNP; delegado(s) de la(s) dirección(es) Técnica(s) del DNP; y delegado de la entidad evaluada, los cuales conforman el Comité de Diseño. En la situación de ejecución de la evaluación hay cuatro posiciones, las mismas tres anteriores conformando esta vez el Comité de Seguimiento, más la posición del contratista/evaluador.

La primera actividad del proceso de diseño de la evaluación[38], es identificar los involucrados en donde se establecen las entidades públicas, las instituciones políticas y sociales claves en el desarrollo de la evaluación, también se identifican los responsables y se convoca a cada involucrado a una reunión de inicio de la evaluación (Grupo de Planeación-DNP, 2013a, p. 1). Esta actividad está a cargo del coordinador de la evaluación y el director de la DSEPP. En otro de sus documentos técnicos se aclara con respecto a este punto:

> "Una evaluación estará mejor formulada, desarrollada y sus recomendaciones serán debidamente discutidas y utilizadas en el ciclo de la política pública, si se involucra en todo el proceso a los actores interesados. La evaluación debe responder a diferentes expectativas y este análisis permite catalogarlas, considerando los intereses de los diversos agentes" (DSEPP-DNP, 2012a, p. 12).

Adicionalmente, se considera que la participación de los actores involucrados aumenta la efectividad y la legitimidad de la evaluación (2014a, p. 34). Todo esto muestra que existen reglas claras sobre la participación de los actores relevantes en etapas tempranas de la evaluación. El alcance de esta participación en la etapa de diseño abarca la elaboración conjunta de la cadena de valor de la política pública; la identificación de problemas y necesidad asociadas a la cadena de valor; y la definición del alcance de la evaluación, que incluye las preguntas a responder, el tipo de evaluación a realizar, y la metodología. Esto significa que existe marco normativo para los estándares de participación de los actores relevantes; consideración de evaluación conjuntas; enfoque asociativo; y estructuras de gobierno y de gestión.

Únicamente hablando de reglas de posición, en la normatividad técnica se resalta la conexión que existe entre un equipo de personas externa a la organización que se evalúa y el hecho de poder generar una mirada objetiva e independiente sobre la intervención Pública (2012a, p. 133). En el *Análisis del sector de consultoría en evaluación de políticas públicas* realizado por la DSEPP, el cual es requisito del proceso de contratación, se establece acerca de la independencia de los evaluadores:

> "… se trata de organizaciones que proporcionan asesoría en los temas o sectores en los que posee un conocimiento especializado. Dentro de las características más relevantes que tienen las firmas que se dedican a realizar consultorías es la independencia, ya que no forman parte integrante de la

[38] Todo el proceso de diseño de la evaluación se encuentra detallado en el documento del subproceso Diseño de la Evaluación, con Código: SE-EP-DE (cuarta versión). Realizado por el Grupo de Planeación del DNP.

> organización que las contrata, razón por la cual pueden ofrecer un punto de vista imparcial sobre el asunto" (DSEPP-DNP, 2014b, p.11).

Que la evaluación sea realizada por una firma consultora externa a la organización evaluada se resalta como requisito para una mirada objetiva e imparcial, por lo que se puede deducir que se tiene claridad acerca de la independencia de los evaluadores. En efecto, siempre se busca contratar organizaciones o equipos de evaluación externos a las entidades evaluadas, por lo que puede considerarse que existe normatividad orientada a satisfacer el estándar de independencia de los evaluadores.

2.1.3 Reglas de elección

En términos generales, las entidades sectoriales evaluadas y las direcciones técnicas del DNP tienen las mismas funciones (Decreto 1082 de 2015, art. 2.2.7.3.1.4.), entre estas están: Apoyar discusiones en la fase de diseño; acompañar la ejecución de la evaluación; proporcionar la información requerida; generar espacios de socialización de los resultados y las recomendaciones de la evaluación; y propender por el uso de estos últimos como insumo para la toma de decisiones.

La DSEPP tiene a su cargo formular documentos de política en los temas de evaluación y seguimiento; establecer lineamientos de evaluación de resultados de la gestión de intervenciones públicas; desarrollo e implementación de mecanismos de evaluación; desarrollo, monitoreo, y evaluación del uso y actualizar los instrumentos de evaluación del Sistema Nacional de Evaluación de Resultados; velar por una adecuada divulgación de las evaluaciones (Decreto 2189 de 2017, art. 15); la coordinación del proceso de evaluaciones, que incluye acompañar desde el punto de vista técnico el diseño y la ejecución de las evaluaciones (Decreto 1082 de 2015, art. 2.2.7.3.2.13.).

Estos tres actores (DNP-DSEPP, DNP-direcciones técnicas y entidades sectoriales), a través del Comité de Diseño, tienen a su cargo el diseño de la evaluación en correspondencia con el proceso de evaluaciones estratégicas estandarizado[39] (art. 2.2.7.3.2.6.). Desde el Comité Técnico de Seguimiento tienen como competencias: apoyar el seguimiento técnico a los productos de las evaluaciones; emitir recomendaciones sobre la calidad técnica de los productos; recomendar parámetros de calidad (Decreto 1290 de 2014, art. 33).

La Secretaría Técnica del Comité Técnico de Seguimiento (a cargo de funcionario/contratista de la DSEPP), tiene a su cargo citar a las reuniones; llevar

[39] Estandarizado en el Sistema de Gestión de Calidad del DNP, y reglamentado en el Decreto 1082 de 2015 (art. 2.2.7.3.2.1.) y el Decreto 1290 de 2014 (art. 25 al art.31).

las actas de estas; distribuir los productos de la evaluación; recibir comentarios a los productos, consolidarlos, depurarlos y darles tratamiento (retroalimentar al evaluador para que corrija los productos de acuerdo con estos comentarios); mantener informados a los integrantes del Comité Técnico de Seguimiento sobre el desarrollo de la evaluación (art. 34).

Estas competencias son desarrolladas a lo largo del proceso de evaluaciones estratégicas, el cual tiene por fases la selección, el diseño, la ejecución y la socialización y uso de la evaluación (art.26). En la fase de selección de la política a evaluar se toman decisiones sobre la priorización de los temas que podrían iniciar proceso de evaluación, a partir de una discusión con todos los sectores de gobierno. Dentro de los criterios que se utilizan para esto, se encuentra la identificación de intervenciones públicas con alta incidencia en los objetivos estratégicos del Plan Nacional de Desarrollo (art. 27).

En la etapa de diseño de la evaluación, el coordinador de la evaluación (funcionario/contratista DSEPP que luego ejercerá la secretaría técnica del Comité de Seguimiento) recopila y analiza información relacionada con la política a evaluar[40], en donde solicita a los responsables de la política documentos y datos (como manual operativo, documentos de crédito, CONPES, documentos de gestión, evaluaciones, etc.) para ampliar y precisar el conocimiento acerca de la intervención (DSEPP-DNP, 2014, p. 38).

Posterior a este análisis y de manera concertada con los involucrados, se construye la cadena de valor (estructura causal de insumos, procesos, productos, resultados e impactos) de la política a evaluar, y se identifican los “cuellos de botella” o restricciones que generan fallas en el flujo de causalidad de esta (DNP, 2016, p. 2 Diseño de la evaluación Código: SE_EP_DE, versión 4; DSEPP-DNP, 2014a, P. 38), a partir de esto se determinan las necesidades de información de las entidades sectoriales que la evaluación abordará.

Se reglamenta que la evaluación tenga en cuenta los planes, actividades y políticas relacionados, así como el análisis de la viabilidad de la evaluación, incluyendo una adecuada definición de la intervención (y sus resultados), y un análisis de la pertinencia con respecto a las necesidades y preguntas de los actores implicados. Por esto, puede decirse que el marco normativo de la política satisface los estándares de evaluabilidad y de coordinación y alineamiento.

La etapa de diseño de la evaluación tiene como elemento central la definición del alcance de la evaluación. Con base en la cadena de valor y en la necesidad de

[40] Este análisis busca responder preguntas acerca del funcionamiento del sector, de los programas con objetivos similares en Colombia, el contexto internacional y las fuentes de información del programa (DSEPP-DNP, 2012b, p. 16).

información identificada anteriormente, se determinan el objetivo general, los objetivos específicos, los componentes, las preguntas orientadoras y los lineamientos metodológicos que se deberán tener en cuenta en la ejecución de la evaluación (DNP, 2016, p. 2 Diseño de la evaluación Código: SE_EP_DE, versión 4). Esto se realiza en el Comité de Diseño, sin embargo, posteriormente la DSEPP hace una revisión de este alcance, en donde existen “ajustes” unilaterales. No se evidencia que se contemple una nueva discusión en el Comité de Diseño de esta revisión unilateral dentro de los lineamientos internos del subproceso de Diseño de la evaluación.

Luego de esto, se consolida todo lo anterior en una presentación con los parámetros y requerimientos para le ejecución de la evaluación, la cual, previa aprobación por parte del nivel directivo (tanto del DNP como de las entidades sectoriales), será el insumo para la construcción de los estudios previos o especificaciones técnicas (conocidos también como Términos de Referencia en la literatura de evaluación) para la contratación del evaluador.

La anterior normatividad jurídica y técnica reglamenta satisfactoriamente todos los componentes de los estándares de objeto y alcance de la evaluación; de objetivos específicos de la evaluación; y de fundamentación y finalidad de la evaluación. De este último, si bien no se encuentran requerimientos puntuales acerca del establecimiento de la utilización prevista, la determinación de las preguntas y objetivos a partir de una demanda de información cubre este punto.

Adicional a lo anterior, para las preguntas de investigación hay guías y lineamientos adicionales en documentos desarrollados por la DSEPP como: Guía de Evaluación de Operaciones; Guía de Evaluación Institucional; Guía de Evaluación de Resultados; Guía de Evaluación de Impacto[41], además de la Guía para la Evaluación de Políticas Públicas (DSEPP-DNP, 2012a) en donde se profunda en mucho mayor grado en el menú de opciones de preguntas y cuestiones a responder para cada tipo de evaluación.

En esta última guía también se direccionan los criterios de pertinencia, eficacia, eficiencia, impacto y sostenibilidad dependiendo del tipo de evaluación y sus preguntas características. Esto significa que los estándares de cuestiones a responder y de selección y aplicación de criterios de evaluación también se encuentran satisfactoriamente reglamentados.

Frente al estándar de selección del enfoque y la metodología, las guías metodológicas mencionadas brindan suficientes lineamientos para garantizar que la metodología se desprenda del propósito, el alcance y las preguntas de la evaluación. Los lineamientos metodológicos generales de cada evaluación quedan

[41] Estas cuatro guías se pueden encontrar en: https://sinergia.dnp.gov.co/Paginas/Internas/Evaluaciones/Metodolog%C3%ADas.aspx

consignados en los estudios previos (que luego se reflejarán en el pliego de condiciones para la contratación), mientras que la metodología de recolección de datos es propuesta por los evaluadores/contratistas, y queda consignada en el primer producto que deben entregar a la DSEPP (tal como se evidenciará más adelante). Así mismo, los criterios SMART de la OCDE para indicadores son retomados por la DSEPP en varias de sus guías metodológicas (DSEPP-DNP, 2012a, p. 67; 2012b, p. 21; 2009, pp. 9-10). En este sentido, el marco normativo incluye todos los componentes del estándar de selección de enfoque y metodología.

Posterior a la definición del alcance de la evaluación, se procede a la definición del cronograma, el recurso humano necesario, y el costeo y el flujo de caja si se realiza en el Comité de Diseño (Grupo de Planeación-DNP, 2016, pp. 2-4). Estos corresponden a los pasos 10, 11 y 12 del proceso de diseño de la evaluación y reglamentan el estándar de realización de la evaluación dentro del plazo y el presupuesto asignados.

Luego de la etapa de diseño, se realiza la gestión de la contratación del evaluador/contratista y comienza la fase o subproceso de “ejecución de la evaluación”. Esta fase inicia con una reunión en donde se define el Comité de Seguimiento, y continúa con la supervisión del desarrollo de la evaluación el cual consiste en “controles permanentes sobre los resultados preliminares y se solicitan las correcciones que se considere necesarias” (Grupo de Planeación-DNP, 2013b, p. 2). La revisión sobre los productos entregados se realiza de conformidad con las condiciones técnicas y su ajuste con respecto a lo que se encuentra en el contrato.

La mayoría de los estándares derivados, o que funcionan conceptualmente como componentes del estándar de informe de la evaluación (explicación de la metodología, lógica de la intervención, validez y fiabilidad de fuentes de información, mención de discrepancias entre miembros del equipo evaluador, Mención de los cambios y limitaciones de la evaluación, contexto de la intervención, y respuesta a las preguntas evaluativas), no cuentan con lineamientos transversales a todos los procesos de evaluación dentro de la normatividad jurídica o técnica, esto es así porque estos lineamientos son determinados para cada contrato en particular al interior del Pliego de Condiciones (términos de referencia) para la contratación del evaluador.

El estándar de claridad del análisis (que también funciona conceptualmente como componente del estándar de informe de la evaluación) se reglamenta a todos los procesos de evaluación al establecer como criterio de revisión de los productos la verificación que los resultados y las recomendaciones, que se detallan en el informe final, estén fundamentados en la información recolectada (Grupo de Planeación-DNP, 2013b, p. 2).

Dentro de los criterios para la revisión del resumen que debe realizar la DSEPP-DNP se establecen el contenido claro y fácil de entender por parte de los tomadores

de decisión y grupos de interés; que incluyan objetivos, resultados más importantes, recomendaciones y una breve descripción de la metodología de la evaluación (2012b, p. 53; Grupo de Planeación-DNP, 2013c, p. 1). En este sentido, de los estándares que funcionan conceptualmente como componentes del estándar de informe de la evaluación, únicamente los de claridad del análisis, y claridad y representatividad del resumen cuentan con marco normativo para todos los procesos de evaluación.

Por último, en cuanto a reglas de elección, las competencias legales de supervisión que tiene el Comité de Seguimiento frente a la elaboración de los productos de la evaluación, en conjunto con las reglas de frontera de todo el proceso de evaluación (que fueron presentadas anteriormente), reglamentan parcialmente el estándar de consultora y protección de los implicados. La protección de los derechos, la confidencialidad y el anonimato de los informantes es algo que se reglamenta al interior de cada contrato.

2.1.4 Reglas de agregación

Las reglas de agregación establecen el nivel de control que los participantes en una posición ejercen en la selección de una acción. Tal como se mencionó en la sección anterior, una de las competencias que tiene la Secretaría Técnica del Comité Técnico de Seguimiento es “Recibir los comentarios a los informes o productos de la evaluación que hagan los otros miembros del Comité Técnico de Seguimiento, consolidarlos, depurarlos y darles el tratamiento respectivo” (Decreto 1290 de 2014, art. 34). Esta competencia se ejerce en correspondencia con los pasos de supervisión del desarrollo de la evaluación y, revisión de los productos consignados en el proceso de ejecución de la evaluación.

Posterior a la revisión conjunta que hace el Comité de Seguimiento, se procede a retroalimentar al evaluador los comentarios y observaciones que se hicieron sobre el producto entregado (Grupo de Planeación-DNP, 2013b, p. 2). Así se procede en esta retroalimentación:

> “El coordinador de la evaluación, como secretario técnico del comité, será el encargado de solicitar a la firma consultora la inclusión de los ajustes en sus productos y justificar técnicamente los requerimientos del comité (…) En caso que se presente una cantidad considerable de comentarios o se identifique la necesidad de hacer ajustes profundos, el coordinador de la evaluación puede convocar al consultor para discutir dichas observaciones (…) en el transcurso de la reunión se hace una revisión y análisis de las recomendaciones y comentarios. Se le expone a la firma consultora los elementos generales, específicos y de forma que deben ser ajustados en el texto.” (DSEPP-DNP, 2014a, pp. 45-46).

Esto también se apoya con un particular Formato de Respuesta a Comentarios[42], el cual corresponde a una matriz de Excel, en esta se ubican como columnas la página en el producto donde se encuentra el comentario, el comentario, la entidad que realizó el comentario, la respuesta de la firma consultora, el estado del comentario y las observaciones. Con el insumo de los productos de la evaluación ajustado, el coordinador de la evaluación procede a realizar un concepto técnico de estos productos, con lo cual se aprueban los productos de la evaluación y se procede a autorizar el pago al evaluador/consultor.

El marco normativo incluye todos los componentes de los estándares de control de calidad y de inclusión de observaciones de los implicados. Sin embargo, debido al fuerte control que se permite sobre los productos y las recomendaciones de los evaluadores, estos dos estándares pueden entrar fácilmente en conflicto con el estándar de proceso evaluativo libre y abierto, particularmente con la independencia del evaluador. La manera en que se plasme esta tensión en la realidad se verá reflejado en el estudio de la coordinación del proceso de evaluación con el de política pública.

2.1.5 Reglas de información

Las reglas de información afectan la información del conocimiento contingente de los actores. En general, el Comité de Diseño y el Comité de Seguimiento, tal como se observó con las competencias presentadas en las reglas de elección, son espacios que permiten que los distintos actores tengan un constante flujo de información acerca de todo lo que sucede en la evaluación.

Frente al estándar de documentación relativa a la finalidad y la expectativa, luego de terminar el proceso interno de diseño, y teniendo como insumo la presentación del alcance de la evaluación, inicia la elaboración de los estudios previos. En el Manual de Contratación del DNP (DNP, 2014a, pp. 6-9) se establecen que los Estudios Previos deben contener:

- Descripción de la necesidad que la entidad estatal pretende satisfacer con la contratación.
- El objeto a contratar, con sus especificaciones, las autorizaciones, permisos y licencias requeridos para su ejecución, y cuando el contrato incluye diseño y construcción, los documentos técnicos para el desarrollo del proyecto.

[42] Se puede encontrar en: https://sinergia.dnp.gov.co/Paginas/Internas/Evaluaciones/Protocolos_y_formatos.aspx

- La modalidad de selección del contratista y su justificación, incluyendo los fundamentos jurídicos que soportan su elección.
- Los criterios para seleccionar la oferta más favorable.
- El análisis del riesgo y la forma de mitigarlo.
- Las garantías que la entidad estatal contempla exigir en el proceso de contratación.

Con base en los estudios previos se construye el pliego de condiciones por parte del Grupo de Contratación del DNP (p. 13), en este, se indica el objeto a contratar, las obligaciones de las partes, el monto asignado para la contratación, los requisitos jurídicos, técnicos, financieros y de calificación de las propuestas. Este documento se pública para darle divulgación al proceso de contratación ante los interesados de participar. Con esto, todos los componentes del estándar de documentación de finalidad y expectativas se incluyen en la reglamentación general. La elaboración del pliego de condiciones, tanto su proyecto como su versión definitiva constituyen el resultado potencial de la situación de acción de diseño de la evaluación.

La última fase (subproceso del proceso de evaluaciones estratégicas) es la socialización y uso de la evaluación. Este subproceso (Grupo de Planeación-DNP, 2013c) tiene como actividades: revisar la ficha resumen y la presentación para exponer la evaluación (productos del evaluador contratista); coordinar la socialización de los resultados de la evaluación; acompañar la identificación de las recomendaciones que la entidad podría implementar; publicar y socializar los resultados de la evaluación. En el Decreto 1290 de 2014 se reglamenta la socialización de la evaluación de la siguiente manera:

> “Una vez se haya aprobado por parte del Departamento Nacional de Planeación (DNP) el producto con los resultados y recomendaciones de la evaluación, se deberá coordinar su presentación a las direcciones técnicas del Departamento Nacional de Planeación (DNP) y las entidades involucradas que sean definidas por el Comité Técnico de Seguimiento y la entidad ejecutora de la intervención evaluada (…) El Departamento Nacional de Planeación (DNP) enviará los resultados y recomendaciones de la evaluación a las autoridades competentes, a efectos de que sirvan como insumo en los procesos de toma de decisiones en temas presupuestales” (art. 35).

Es la DSEPP quien realiza todo el proceso de socialización de la evaluación, no se encontró evidencia dentro de la reglamentación general de que se incluya la participación del evaluador contratista en este proceso. Aun así, todos los componentes del estándar de difusión (formato accesible, difusión sistemática, difusión externa e interna) se encuentran incluidos en esta reglamentación.

2.2 Coordinación entre el proceso de evaluación con el de política pública

A continuación, se presentan los resultados del análisis de los distintos componentes que teóricamente se delimitaron para el abordaje de la coordinación del proceso de evaluación con el de política pública. Esto se realiza utilizando datos provenientes de la institucionalidad de la política y de entrevistas a funcionarios de la DSEPP y a diversas firmas de consultorías que han sido evaluadores dentro de la política en cuestión.

Si bien, muchos de los elementos seleccionados del U-FE para la operacionalización de la coordinación como adaptación mutua pueden ser ubicados dentro de los tipos de reglas señalados anteriormente, al remitir esta coordinación a la variable "patrones de interacción" dentro del IAD de Ostrom, se prefirió elegir como estructura de análisis y del texto, directamente las variables seleccionadas del U-FE de Patton. Como podrá observarse, analizar la pertinencia de las variables seleccionadas como componentes de la coordinación, necesita ir más allá de la institucionalidad, y adentrarse en el modo en que los participantes actúan en cooperación. Por esta razón, este análisis requirió también el uso de entrevistas.

2.2.1 Involucramiento de los usuarios previstos

Dentro de los componentes de la coordinación como ajuste mutuo se encuentra hallar e involucrar a los usuarios previstos en todo el proceso de evaluación. Además de tener institucionalización amplia y desarrollada, este involucramiento también es ampliamente reconocido por los funcionarios del DNP como el requerimiento más importante para que las evaluaciones sean utilizadas:

> "… la recepción de la recomendación depende mucho del involucramiento que haya tenido el Comité durante todo el proceso (…) que se vea que están como con la disposición y como con la actitud, y que realmente la necesidad, porque si esto pasa durante todo el proceso, también vamos a garantizar un buen diseño, y además en la etapa final vamos a garantizar que las recomendaciones sean acogidas de buena forma" (Entrevista coordinadora grupo de evaluaciones estratégicas de la DSEPP).

> "… es fundamental, muy fundamental que la entidad que es la ejecutora de la política esté involucrada en todo el proceso de la evaluación. Es decir, desde que se proponen los temas de la agenda de evaluaciones. Si la entidad no está involucrada desde el principio, sino que uno le llega al final con los resultados, es un tema en el que de pronto ellos no tenían el interés, les

> estamos llegando con un documento en un momento en que tal vez tienen otras prioridades" (Entrevista contratista DSEPP 1).

En general, existe un gran interés y es fácil constatar que las entidades sectoriales son involucradas desde el inicio en el proceso de evaluación. El problema, en este particular aspecto de la coordinación, es poder sostener la participación de los funcionarios de las entidades sectoriales en los comités descritos en el apartado anterior, esto se debe principalmente al alto nivel de rotación de personal que existe en las entidades públicas en Colombia:

> "¿Qué nos ha pasado? Que las entidades cambian de… pues como hay tanta rotación de personal en el sector público, generalmente cambian los comités, y entonces cuando llega un nuevo comité que no estaba tan involucrado desde el principio, se puede ver un poco el riesgo de eso, y nosotros, es un trabajo de SINERGIA hacer una sensibilización de la cultura de la evaluación, enseñarles porque es importante esta evaluación, entonces tratamos de hacer eso" (Entrevista coordinadora grupo de evaluaciones estratégicas de la DSEPP).

Sin embargo, iniciar con nuevos participantes -ya sea en el Comité de Diseño o en el Comité de Seguimiento el proceso de evaluación- implica invertir tiempo y recursos institucionales adicionales que fácilmente pueden significar obstáculos o retrasos en el desarrollo de la evaluación y, por ende, riesgos en la relevancia temporal de los resultados de la evaluación, lo cual se tendrá que sopesar con la relevancia que los nuevos funcionarios les asignen a estos resultados. En una firma consultora resaltan los problemas que pueden surgir al respecto:

> "No había una entidad clara que hubiera contratado (solicitado) la evaluación. Entonces nunca supimos bien quien la encargó. Creemos que surgió por iniciativa de alguien en el DNP. Pero la Unidad de Víctimas que era la entidad encargada de implementar la reparación colectiva dijo 'aquí no fue, y gracias, pero en realidad pues no lo estábamos esperando' Y fue una evaluación costosa." (Entrevista consultor del Centro Nacional de Consultoría).

Tal como se observa, el papel del DNP puede llegar a ser ambiguo, puesto que puede tener funciones claras del evaluador propuestas por Patton (como apoyo y asesoramiento metodológico en el Comité de Diseño, además de la identificación e involucramiento de usuarios), tener funciones de solicitante/cliente de la evaluación (como se observa en la cita), y tener la función de supervisor de la evaluación, en donde viene a ser el eje central de comunicación para la coordinación entre usuarios previstos y evaluador. Esto es un riesgo para la clara identificación de los usuarios de la evaluación, así como para la comunicación fluida que requiere la coordinación evaluador-usuarios previstos.

Otro de los fallos de los cuales se encuentra información, y que se relacionan con hallar e involucrar a los usuarios de la información, consiste en la ausencia de otros actores implementadores de la política pública evaluada dentro del proceso de

evaluación, en algunos casos se manifiesta ausencia de entidades de otros niveles de gobierno (Entrevista consultor Infométrika), y -en otros casos- se evidencia ausencia de otros sectores del nivel nacional:

> "Como que precisamente es un sistema que funciona cuando se articulan una cantidad de entidades del Estado cada una con un rol diferente. Y a la hora de la verdad ese sistema no es de nadie, y por lo tanto no hubo nadie que se hiciera cargo como tal de ella, y tampoco estaba claro quien había encargado la evaluación. Entonces a veces sí pasa que la misma entidad la encarga, pero a veces pasa, como en estos ejemplos que mencioné, que no es tan claro quién es la contraparte en la entidad que pide la información o que se va a beneficiar de ella" (Entrevista consultora Centro Nacional de Consultoría).

Esto refleja que si bien institucionalmente existen reglas claras para que tanto el Comité de Diseño y el Comité de Seguimiento incluyan delegados de las entidades sectoriales evaluadas, y que en la medida de lo posible la DSEPP se esfuerza por conservar la continuidad de los mismos funcionarios a lo largo de todo el proceso de evaluación, esto no quiere decir que, en efecto, estos funcionarios correspondan a un conjunto satisfactorio de usuarios primarios de la evaluación, ni que tengan la posibilidad de implementar las recomendaciones producidas.

2.2.2 Equipo de trabajo conjunto

Al hacer preguntas puntuales acerca de la relación que se establece entre los actores públicos y el evaluador/contratista, para el establecimiento del equipo de trabajo conjunto entre evaluadores y usuarios, en correspondencia con la institucionalidad analizada (particularmente en las reglas de agregación) se contesta:

> "… la relación que se establece es una relación de supervisión, donde este comité técnico revisa los productos y da recomendaciones y parámetros para que estos productos tengan la calidad técnica en temas de la experticia y del área" (Entrevista coordinadora grupo de evaluaciones estratégicas de la DSEPP).

> "Ahí, lo que pasa es que la relación no es tan directa, porque SINERGIA, y nosotros como DSEPP, somos como los que estamos ahí en la mitad (…) ellos no tienen un contacto directo, siempre es por intermedio del comité técnico de seguimiento y de nosotros como DSEPP. Entonces, las relaciones, la fluidez depende de cómo nosotros manejemos esa relación" (Entrevista contratista DSEPP 1).

> "… es una relación en la que las firmas se entienden con el comité técnico de seguimiento a través de las reuniones periódicas que se realizan para el seguimiento al desarrollo de la evaluación, pero siempre se da una relación a través de nosotros que cumplimos la función de secretaria técnica" (Entrevista contratista DSEPP 2).

Este papel como intermediario también se refleja en que los consultores no pueden pedir directamente información a la entidad evaluada, sino que deben primero hacer la solicitud a la DSEPP para que está la comunique a la entidad (Entrevista contratista DSEPP 1). Esto, sumado al rol ambiguo que tiene el DNP, significa que en términos teóricos se puede hablar de un elevado grado de control por parte de la DSEPP en la relación evaluador-usuario previsto, que incide en la fluidez de los ajustes mutuos que se deben dar entre estos.

Al respecto, se debe precisar que los contratistas de la DSEPP aclararan que en el momento en que ellos retroalimentan al consultor con los comentarios de los funcionarios de las entidades sectoriales, dependiendo de la complejidad de los comentarios, así como de la respuesta de los consultores, se puede citar o no una reunión que será intermediada por la DSEPP (Entrevista contratista DSEPP 1).

Es importante aclarar, por un lado, que no es total el control que ejerce la DSEPP y las entidades sectoriales sobre los resultados y las recomendaciones desarrolladas por el consultor. Se conserva cierto grado de independencia del evaluador únicamente en la posibilidad de defender técnicamente sus conclusiones y recomendaciones, siempre y cuando estos se encuentren dentro de los lineamientos de los términos de referencia.

Por otro lado, las entrevistas también reflejan problemas en esta dinámica de trabajo conjunto. Particularmente se evidencian falencias en el ajuste que debería realizar la DSEPP en orden de coordinarse con las entidades sectoriales debido a su elevado grado de control:

> "… el DNP en general es una entidad que tiene ese problema, le es muy difícil hablarle al resto de las entidades del sector público en otro lenguaje. Tienen sienta arrogancia técnica que les hace estar muy lejos de esas entidades. Y en el caso de las evaluaciones, por ellos tomar el control sucede lo mismo; se crea una barrera de comunicación con el resto de entidades" (Entrevista consultora del Centro Nacional de Consultoría).

Por esta razón, esta consultora defiende una profundización de la relación (y en la comunicación) evaluador-entidad sectorial, a partir de la sugerencia de otro grupo de trabajo "…dentro de las entidades, en el cual se hiciera un esfuerzo mayor por aterrizar los resultados a que todo el mundo lo pueda entender" (Ibíd.). Un consultor de otra organización también problematiza esta relación DNP-entidad sectorial, haciendo énfasis en que es más fácil que las entidades acepten los resultados de la evaluación cuando ellos mismos contratan a la firma consultora, que cuando se

debe establecer el diálogo a través del DNP (Entrevista consultor Universidad de Los Andes).

Este elevado grado de control es perjudicial para la dinámica necesaria que debe existir en el grupo de trabajo conjunto, en donde debería facilitarse la voluntad de trabajo para negociar temas clave que afecten la credibilidad y el uso de la evaluación (Patton, 2013, p. 4). Esta negociación implica necesariamente ser reactivo al escuchar a los usuarios previstos y en responder a sus necesidades y demandas.

2.2.3 Reacción a las demandas, preocupaciones e inquietudes de los usuarios previstos para la construcción conjunta de preguntas de evaluación.

El elevado control también impacta negativamente al escuchar las preocupaciones y las demandas de los usuarios previstos. En este punto, el evaluador debe sintonizarse con las preocupaciones de los usuarios y no permitir que el interés del evaluador (en este caso, el mismo DNP), domine o controle el ajuste de prioridades (2013, p. 9). En la coordinación que se da, este asunto se encuentra íntimamente relacionado con la revisión de opciones de preguntas de la evaluación, en cuanto a esto la DSEPP manifiesta:

> "… los convocamos a ellos, para que ellos nos digan que es lo que quieren y con base en todo eso que ellos quieren formulamos los objetivos de la evaluación y las preguntas de investigación que queremos responder, también es importante la guía de ellos, porque ellos conocen la información disponible y como base en ese conocimiento que ellos tienen de información disponible, podemos formular las preguntas que se puedan responder con base en la información que existe" (Entrevista coordinadora del grupo de evaluaciones estratégicas de la DSEPP).

> "… empiezan a decirnos dónde están los cuellos de botella para esa política, y uno empieza a decir que la entidad lo que necesita es mejorar el proceso, que el impacto o el resultado en el beneficiario se den de una mejor manera, y ahí se empiezan a desencadenar las preguntas orientadoras que requiere la evaluación. Pero, en el diseño, eso es lo que se hace, y ahí uno se da cuenta del alcance real de la evaluación, para saber a dónde apuntarle" (Entrevista contratista DSEPP 1).

Ellos manifiestan cumplir el proceso que está institucionalizado, y tienen una percepción de su control sobre el proceso de construcción de las preguntas, simplemente como una asesoría que se realiza acerca de las necesidades de las entidades sectoriales que debe ser sopesada con la disponibilidad de información.

Sin embargo, la relevancia de las preguntas que resultan de este proceso tiene otra perspectiva desde el lado de los evaluadores/consultores:

> "…realmente el alcance es muy amplio. Es decir, las preguntas orientadoras son terriblemente ambiciosas, y precisamente por eso es que las evaluaciones terminan complejizándose tanto" (Entrevista consultora del Centro Nacional de Consultoría).

> "… muchas veces los funcionarios, de buena fe, dicen que eso está consignado en las bases de datos, que piensan que son buenísimas. Pero cuando uno ve al detalle, esa no es la información que se necesita, o la requerida para responder las preguntas que ustedes mismos nos están haciendo. Entonces, bueno, uno tiene que darse la maña de poder responderla" (Entrevista consultor de Econometría).

> "… cuando generan las preguntas de investigación (…) todas las entidades tienden a ser exhaustivas y a pedir respuestas sobre un montón de cosas que, de pronto, no son tan fáciles de responder con el tamaño de muestra o con los instrumentos que se presupuestan. Entonces muchas veces algunas preguntas de investigación, que ellos las llaman preguntas orientadoras, le complican a uno un montón la vida" (Entrevista consultor de Econometría).

Lo ideal sería que el balance conjunto que se realice entre información requerida y factibilidad de respuesta con la información disponible se haga en el marco de las negociaciones entre el evaluador y los usuarios previstos. El evaluador debe ser quien guie a los usuarios en el entendimiento de que tipos de preguntas pueden ser contestadas con los datos disponibles (Patton, 2013, p. 9). Sin embargo, esto es realizado por parte de la DSEPP (otra función de evaluador ejercida por el DNP), haciendo que el evaluador no pueda manifestarse con respecto a la factibilidad de responder las preguntas de evaluación.

En este sentido, las falencias que puedan existir en la coordinación, particularmente al escuchar y tener en cuenta las preocupaciones y demandas de los usuarios, y al revisar conjuntamente las opciones sobre preguntas de evaluación, se relacionan con la ausencia de participación del contratista en las etapas de diseño de la evaluación, y con las limitaciones en la comunicación que existen entre el evaluador/contratista y los usuarios primarios.

En todo caso, no puede decirse que siempre suceda lo anterior (lo cual deberá ser observado en los casos específicos), puesto que se ha manifestado la existencia de evaluaciones en las que este problema ha podido ser direccionado gracias al dialogo directo con el evaluador/contratista: "… cuando se ha establecido un diálogo técnico fluido, esas cosas se pueden discutir y se le puede dar una interpretación adecuada, y a veces incluso incluir nuevas preguntas que sí se pueden responder y que resultan más interesantes, y son más relevantes para las evaluaciones" (Entrevista consultor de Econometría).

2.2.4 Trabajo conjunto sobre objetivos prioritarios y usos previstos

Como se observó en el análisis institucional, en la etapa de diseño de la evaluación existe la actividad conjunta entre DSEPP y entidades sectoriales para identificar los problemas y las necesidades asociadas a la cadena de valor, la cual continúa con la definición conjunta del alcance de la evaluación, en donde se definen preguntas y objetivos. De esta discusión, en el Comité de Diseño se desprende la determinación de los usos de la evaluación:

> "Lo importante cuando uno está en el diseño de la evaluación es construir bien la cadena de valor, los insumos los procesos, los productos y los resultados, y los cuellos de botella que hay en cada uno de esos eslabones de la cadena. Ahí es donde uno realmente se da cuenta hacia dónde debería apuntarle la evaluación" (Entrevista contratista DSEPP 2).

Como se observa, en la etapa de diseño no hay una discusión puntual acerca de los usos a los cuales se va a dirigir la evaluación, por lo menos no en los términos de Patton, sin embargo, lo que está implícito es que la información que se obtenga, así como los hallazgos y recomendaciones, se usen para solucionar los denominados "cuellos de botella o restricciones que generan fallas en el encadenamiento necesario para alcanzar los resultados esperados" (Grupo de Planeación-DNP, 2016, p. 2 Diseño de la evaluación Código: SE_EP_DE, versión 4).

El evaluador/contratista no participa en esta discusión al ser una actividad previa a su contratación, por esto, sólo se podría hablar de un ajuste a partir de la adaptación de este participante a lo que ya ha sido establecido por el Comité de Diseño. Es posible que la imposibilidad de que el consultor/evaluador de mostrar las distintas opciones de usos, o proponer usos particulares de acuerdo con su valoración del contexto particular de la intervención, en las etapas tempranas del proceso de evaluación, impacte negativamente en el uso de las recomendaciones y hallazgos que ellos mismos construyen.

2.2.5 Debates metodológicos entre los usuarios previstos y el equipo evaluador

Al pertenecer la selección del enfoque metodológico y la construcción de las preguntas de evaluación a la misma actividad de definición del alcance de la evaluación, las cuales son previas a la contratación del evaluador, un fenómeno similar sucede con las discusiones metodológicas que integran la coordinación estudiada, pero no con la misma profundidad. Por su parte, en la DSEPP se resalta que en los últimos años se le ha ido dando más peso a la propuesta metodológica

de los postulantes a ser evaluadores/contratistas como criterio de calificación: "… para calificar y escoger al mejor, y que se haga la mejor evaluación, es darle más peso a lo técnico y a lo metodológico. Entonces, se creó un sistema en el que damos puntajes, dependiendo el tipo de evaluación" (Entrevista contratista DSEPP 1).

Entonces, existen unos requisitos metodológicos ya establecidos en el pliego de condiciones para la contratación del evaluador. Los postulantes deben presentar una propuesta metodológica que desarrolle los requisitos metodológicos del pliego de condiciones los cuales no son negociables. Al preguntar sobre este punto, la DSEPP responde que los postulantes centran su propuesta metodológica al momento de la postulación en la metodología de recolección de datos, y que, en todo caso, en etapas posteriores pueden ir más allá de lo que se les exige en el pliego de condiciones:

> "… eso es lo que ellos están obligados a hacer, pero ellos en su experiencia en investigación pueden llegar a responder otras cosas, o una misma pregunta orientadora pero más especifica que tenga en cuenta otras variables que surgieron en diseño, pero que luego del levantamiento de la información y del análisis de la información secundaria, se dieron cuenta de que pudieron responder esas cosas, lo puede incluir, pero mínimo las preguntas orientadoras que nosotros establecemos" (Entrevista contratista DSEPP 2).

Adicional a esto, existe un espacio en el que estos requisitos metodológicos pueden modificarse, se trata de la audiencia de aclaraciones (Grupo de Planeación-DNP, 2010; y Decreto 1082 de 2015 en su artículo 2.2.1.1.1.7.1.), la cual se ubica entre el momento en que el Grupo de Contratación del DNP saca un proyecto de pliego de condiciones, lo da a conocer y los postulantes pueden presentar preguntas y recomendaciones, los cuales son estudiados y seleccionados o descartados por parte del DNP para que aparezcan el pliego de condiciones definitivo para la contratación.

Se deberá hacer una comparación al interior de cada caso del documento de proyecto de pliego de condiciones y su versión definitiva, para caracterizar el grado en el que las propuestas metodológicas de los consultores son tomadas en cuenta. Sin embargo, aunque esto se vea reflejado, la discusión metodológica en este punto se está dando entre el consultor y el DNP y no entre el evaluador y los usuarios previstos, lo cual limita la coordinación entre los procesos.

2.2.6 Adaptación conjunta del diseño de la evaluación

La discusión que se realice acerca de la metodología de la evaluación no se puede reducir a las etapas previas a la ejecución de la evaluación. Parte importante de la coordinación como ajuste mutuo para que una evaluación se use, son las

discusiones que se deben llevar a cabo para la adaptación de la metodología a medida que la evaluación se va desarrollando (Patton, 2013, p. 11). Esta adaptación es necesaria porque las evaluaciones pueden llegar a demorarse entre dos años o dos años y medio (Entrevista consultor Centro Nacional de Consultoría; Entrevista consultora Centro Nacional de Consultoría) desde el momento en que la entidad sectorial manifiesta su demanda de información, hasta el momento en que tiene los últimos productos de la consultoría.

En este tiempo, es muy probable que varios de los aspectos de la evaluación cambien, como por ejemplo la demanda de información o la factibilidad de responder a las preguntas orientadoras[43], razón por la que los consultores han identificado que el empeño del DNP en asegurar objetividad y rigurosidad puede significar un sacrificio en la relevancia y en el impacto que pueda tener la evaluación en una entidad (Entrevista consultor Universidad de los Andes). Caracterizando más a profundidad este fenómeno, se manifiesta:

> "… se enfrasca en un tema de representatividad [de la muestra], porque muchas veces ahí es donde los proyectos si bien ganan en un tema académico (…) desde el punto de vista de política, son proyectos que pierden vigencia y relevancia muy rápido. Los casos en los que no ha sido así son contados con la mano. Se me ocurre por ejemplo Familias en Acción (…) que se convirtió en una política de Estado (…) Pero en el resto de los programas uno ve que son de corta duración y que las evaluaciones duran más de lo que esto" (Entrevista consultor Centro Nacional de Consultoría).

> "Yo creo que el diseño de las evaluaciones que solicita el DNP son más evaluaciones para ámbitos académicos que para ámbitos políticos. Entonces, ¿qué sucede la mayoría de las veces? Son evaluaciones sofisticadas, con una complejidad grande, que además toman mucho tiempo, y, finalmente, cuando salen los resultados, las prioridades de política ya han cambiado, los funcionarios ya han cambiado, todo ha cambiado. Nos ha pasado incluso que antes de que salgan los resultados de la evaluación el programa ya lo han terminado porque no es una prioridad del funcionario que está de turno. Por todo esto, pienso que finalmente debería haber un balance entre la complejidad técnica y los tiempos que se requieren y la pertinencia del momento o la oportunidad de la evaluación" (Entrevista consultor Centro Nacional de Consultoría).

[43] Tal como se observó cuando los consultores manifiestan que la percepción que los funcionarios de sus bases de datos o su información, no siempre corresponde a las posibilidades reales que esta tiene para responder a las preguntas que ellos mismos formulan (Entrevista consultor de Econometría).

Este balance entre complejidad (rigurosidad) técnica y la oportunidad de la evaluación, no es una discusión que deba limitarse a la etapa de diseño de la evaluación, se requiere de una coordinación y un diálogo continuo con los usuarios previstos para que los evaluadores se puedan adaptar al cambio de su demanda. Esto no se realiza, tampoco se encuentra contemplado en la institucionalidad de la política, y es probable que el papel de intermediario del DNP en el dialogo entre evaluador y usuarios previstos, así como su elevado control de todo el proceso[44] resulten ser obstáculos para la adaptación de la evaluación.

Puede hablarse de una inflexibilidad o rigidez del diseño de la evaluación, que también es reforzada por las exigencias y lineamientos técnicos que se consignan en los contratos de consultoría, y por el temor a las sanciones que llevaría el incumplimiento de sus cláusulas:

> "… por el miedo que se tiene a alguna investigación por parte de un organismo de control para ver si realmente sí se estaba cumpliendo al pie de la letra lo que se contrató, hace que el desgaste en el desarrollo de las evaluaciones sea muy grande. Es decir, casi que a veces no se toman decisiones con sentido común técnico sino con sentido jurídico dependiendo de lo que el abogado autorice o no autorice. Para las firmas eso es terriblemente desgastante, y para el mismo Comité de Seguimiento a la evaluación y todas las entidades también" (Entrevista consultora del Centro Nacional de Consultoría).

Esta rigidez en el diseño de la evaluación termina redundando en una limitación para la apropiación de la evaluación por parte de los usuarios primarios:

> "… hay una falla, y es que como viene la solicitud de términos de referencia de las evaluaciones de DNP en el diseño de las mismas, por esa exigencia de calidad técnica es que a veces se genera un exceso de tecnicismo. Esto hace que las evaluaciones no sean aterrizadas al lenguaje de los funcionarios de estas entidades y finalmente por eso se hace más difícil la apropiación también" (Entrevista consultora Centro Nacional de Consultoría).

En términos teóricos se puede decir que no existe un ajuste mutuo para la adaptación del diseño y de la metodología de la evaluación a las necesidades cambiantes que tienen los usuarios previstos. Esto significa una fuerte limitación a la posibilidad que tiene el equipo de trabajo conjunto (Comité de Seguimiento) a responder a la demanda de la información y, por ende, a que esta se use.

---

[44] Los consultores perciben que al DNP "…le cuesta mucho soltar ese control que tiene de la evaluación" (Entrevista consultora Centro Nacional de Consultoría).

2.2.7 Trabajo conjunto en la interpretación de hallazgos y la construcción de recomendaciones

En cuanto a las discusiones y ajuste mutuo que se debe dar en el involucramiento de los usuarios previstos en la interpretación de los hallazgos y en la construcción de las recomendaciones de la evaluación, además del marco institucional en que se brinda la oportunidad de comentar todos los productos, la DSEPP reconoce que son los diseñadores e implementadores de las políticas quienes tienen el conocimiento necesario para dar parámetros acerca de la factibilidad de las recomendaciones, brindar valoraciones acerca de la facilidad en la implementación de estas, y determinar a los actores que serían responsables de implementarlas. Así es caracterizado esto por parte de un consultor:

> "… cada producto que uno presenta, tanto el metodológico, como el avance, como las conclusiones, como las recomendaciones, las [entidades sectoriales] comentan. Y el DNP concilia esos comentarios para que el consultor haga las correcciones, las aclaraciones, o presenten razones de lo que escribieron, o que ahonden más en lo que les señalen. Y de correcciones, desde cuestiones de redacción hasta cuestiones de forma, de interpretación de los resultados" (Entrevista consultor Econometría).

Fácilmente se puede advertir del riesgo que esto implica para la independencia del juicio del evaluador, sin embargo, los participantes subrayan que esto no afecta la imparcialidad del evaluador/consultor:

> "…aunque el comité siempre orienta mucho la formulación de las recomendaciones, al final la firma es autónoma y por eso es que contratamos externos, porque son terceros imparciales, entonces al final la firma es la que tiene la discreción sobre qué recomendaciones formular, entonces ellos, la firma generalmente es muy receptiva de las recomendaciones de los parámetros técnicos, pero en últimas ellos son autónomos" (Entrevista coordinadora del Grupo de Evaluaciones Estratégicas de la DSEPP).

Al respecto, los consultores no revelan una percepción que pueda ser considerada muy distinta:

> "Nosotros cuidamos mucho que las recomendaciones sean muy equilibradas. Rara vez una evaluación que nosotros hagamos, y estoy casi segura que ninguna firma lo haría, tiende a hacer recomendaciones especialmente controversiales o a ser evaluaciones con un enfoque negativo de acabar con el programa. Todo lo contrario, creo que todos nos cuidamos de tener un enfoque apreciativo, así lo llamamos acá, en donde se destaquen tanto las fortalezas del programa como las cosas que en efecto hay que mejorar" (Entrevista consultora de Centro Nacional de Consultoría).

En efecto, puede considerarse que existe un diálogo y un ajuste mutuo para la interpretación de los hallazgos y la construcción de las recomendaciones. Sin embargo, los consultores también hacen la salvedad de que pueden existir situaciones en que la firma consultora queda en medio de discusiones y rivalidades interinstitucionales con respecto a la implementación de una política determinada que no han sido resueltas (Entrevista consultor de Centro Nacional de Consultoría). Esto termina limitando la posibilidad de apropiación de las recomendaciones por parte de los usuarios previstos.

2.2.8 Planeación conjunta del seguimiento a la implementación de los hallazgos y recomendaciones de la evaluación

El involucramiento de los usuarios en la elaboración de las recomendaciones va de la mano con la construcción de un plan conjunto de seguimiento a la implementación de las recomendaciones. Una construcción conjunta con respecto a este punto se debe ver reflejada al interior del Plan de Transferencia e Implementación de Recomendaciones (PTIR), que es uno de los últimos productos entregados por los evaluadores/consultores.

A pesar de la existencia del PTIR, una vez se consideran aprobados los últimos productos del consultor y se realiza el último pago, existe una clara ruptura dentro del proceso de coordinación, esto es reconocido por la DSEPP:

> "…entonces nosotros tratamos de hacerles seguimiento, ahorita por ejemplo, todavía le hacemos seguimiento a las evaluaciones que terminaron en 2017, pero eso no va a ser más de este año, o más de mediados de este años, ya de ahí para delante no solo no es competencia nuestra, sino que también un poco nos desligamos de los comités, como que se acabó un poco el marco de la evaluación, entonces ya no hay más reuniones, la relación digamos se diluye un poco (…) nosotros hacemos seguimiento durante un tiempo. Por ahí unos 8 meses (…) incluso a veces hasta por un año, pero nosotros realmente la competencia deja de ser nuestra apenas nosotros les entregamos los resultados a las direcciones técnicas del DNP" (Entrevista coordinadora del Grupo de Evaluaciones Estratégicas de la DSEPP).

A parte del hecho de que se debería considerar a los usuarios primarios como los clientes principales de los productos de la evaluación, La comunicación debe ir más allá del reporte final. Patton manifiesta que uno de los más grandes obstáculos para trabajar con los usuarios previstos para aplicar y actuar en los hallazgos y recomendaciones es que los diseños, planes y contratos de evaluación típicamente terminan en el reporte (2012, p. 381), esto significa no asignar tiempo o dinero para el seguimiento que facilite el uso de la evaluación. En efecto, esta es la situación que ocurre con la política de evaluación.

El remplazo de la Matriz de Gestión de Recomendaciones por la Matriz de uso (como se expone en el subtítulo de reglas de alcance), refleja que ese seguimiento que actualmente se realiza es unilateral por parte de la DSEPP, por lo que no constituye un ajuste mutuo entre los participantes y, por consiguiente, no puede ser considerado un escenario de coordinación para el uso de la evaluación.

También es de resaltar que justifique la terminación del seguimiento a partir de la ausencia de competencias desde el momento en que se entregan resultados a las direcciones técnicas del DNP. Esto implica un desconocimiento intencional de la responsabilidad de la DSEPP con respecto a la incidencia en las decisiones de las entidades sectoriales, esto se refleja mejor en la siguiente cita:

> "Nosotros nos encargamos de hacer una evaluación, que nos demandó un sector, como te decía, el PTIR, las recomendaciones no son vinculantes, tratamos de hacer que lleguen a los niveles directivos más altos, tratamos de sistematizarlas, pero no es vinculante, porque además el papel de SINERGIA es evaluar, nosotros no podemos incidir en las decisiones de las demás entidades" (Ibíd.)

Se dice esto, a pesar de que el eslogan de SINERGIA en los últimos años ha sido "incidiendo con evidencia", y que en los decretos que reglamentan a SINERGIA, las reglas de alcance establecen que la evaluación de políticas públicas que realice el DNP debe servir para generar información y evidencia, que sirva como insumo para la toma de decisiones y para mejorar el diseño, la implementación y los efectos de las intervenciones públicas (Decreto 1290 de 2014, art. 21, art. 22 y Decreto 1082 de 2015, Art. 2.2.7.3.1.4.).

Lo anterior es el objetivo central del componente de evaluaciones de SINERGIA, la efectividad (la medida en la que se cumplen los objetivos) en su cumplimiento no puede medirse a partir de la disposición de información (no es una medición válida), sino a partir del grado en que la información y la evidencia se reflejan al interior de la toma de decisiones, para esto, el medio conceptualmente concreto es ver el grado de implementación de las recomendaciones, y es por esta razón que se diseñó e implementó la Matriz de Gestión de Recomendaciones.

2.2.9 Búsqueda proactiva del uso de la evaluación

Por último, con respecto la búsqueda proactiva del uso de la evaluación, esta tampoco se vería reflejada por la ruptura en la comunicación al momento de terminar el contrato de consultoría. En este punto el evaluador debería ajustarse a los usuarios primarios de tal modo que estén atentos a oportunidades para reforzar la relevancia de los hallazgos a partir de comunicar los hallazgos en momentos

específicos[45] (Patton, 2012, p.382). Esto no sucede, e incluso puede caracterizarse como precaria la participación de los evaluadores/consultores en la difusión de los resultados que ellos mismos construyeron:

> "Muchas veces nos dicen, una vez que ya hemos entregado y quedan aprobados todos los documentos, que nos van a pedir que en unos días hagamos una divulgación en cierto espacio, y aunque por supuesto estamos abiertos a hacerla eso nunca pasa. Nunca llega, nos dicen que las agendas están muy complicadas. La idea se va diluyendo en el tiempo y nos quedamos esperando y dispuestos para ir a hacer la presentación y no sucede." (Entrevista consultora del Centro Nacional de Consultoría).

> "… realmente nosotros no tuvimos interacción de los resultados con las entidades. Nosotros hicimos, por lo menos hasta donde Infométrika llegó, la socialización de resultados con el mismo DNP; más allá de ni con entidades territoriales, ni con ministerio. Este último estuvo en algunas reuniones y algunos resultados, pero más allá de eso no podría responder cómo percibieron ellos eso" (Entrevista consultor de Infométrika)

En este sentido, los evaluadores/consultores a pesar de ser los principales proponentes de los hallazgos y las recomendaciones de la evaluación, tienen posibilidades muy limitadas para la búsqueda proactiva del uso de estos, así como frente al proceso de difusión en el cual se enmarcaría. Podría pensarse que el DNP sea quien lidere la búsqueda proactiva del uso de los estudios en los que invierte su presupuesto, lo cual sería equivalente a lo que denominan "incidencia", sin embargo, su postura frente a las competencias en cuanto a la implementación de las recomendaciones descarta esta posibilidad.

2.3 Percepción general acerca de la efectividad de los procesos de evaluación

La efectividad de los procesos de evaluación, así como de toda la política de evaluación en general, debe ser abordada a partir de la caracterización a profundidad del uso de la evaluación, cuya dimensión concreta es el grado en el cual se han implementado las recomendaciones de la evaluación. En esta sección no se realiza esta caracterización en profundidad, sin embargo, se presentan las percepciones que tanto funcionarios de la DSEPP como consultores/evaluadores tienen acerca de esta dimensión concreta. Esto para brindar una primera aproximación a la efectividad general de la Política de Evaluación de Políticas en Colombia.

[45] Como nuevas decisiones, eventos no anticipados, nuevas investigaciones u otra información.

Si bien esta percepción tiene un alcance muy limitado frente a la posibilidad de falsear o comprobar las teorizaciones bajo estudio, permite tener una idea general y exploratoria acerca del uso de las evaluaciones, y de los factores que lo restringen o lo fomentan en este particular contexto. Esto tiene el potencial de ofrecer inferencias que posteriormente pueden ser contrastadas con el análisis comparado. La sección se divide en tres partes, primero una percepción general acerca del uso de las evaluaciones por parte de las entidades sectoriales evaluadas. En segundo lugar, se describen los factores que los entrevistados asocian como explicación de la variación del uso de las evaluaciones.

Y finalmente, se presentan los resultados de la implementación de las recomendaciones realizadas en los procesos de evaluación a SINERGIA en 2010 y 2013 por parte del DNP. Como se verá, en ambas evaluaciones se recomendó al DNP promover el uso de los resultados de las evaluaciones, por lo que este ha sido un criterio de meta-evaluación pertinente en el caso colombiano, por lo menos desde estos años.

2.3.1 Percepción general frente a uso de las evaluaciones

Al realizar preguntas acerca de la recepción de los resultados de las evaluaciones y de la implementación de las recomendaciones a funcionarios de la DSEPP, en general, realizan un balance medianamente positivo.

> "… Cómo lo he visto, en general ellos acogen muy bien esto (…) Si, digamos la competencia de SINERGIA llega hasta cierto punto, pero si he visto que hay evaluaciones, no ya desde el papel de SINERGIA, pero si he visto que en las entidades hay ciertas evaluaciones que se han utilizado" (Entrevista coordinadora del Grupo de Evaluaciones Estratégicas de la DSEPP).

> "Bueno, como el PTIR es una herramienta relativamente nueva que se ha incluido en el proceso de las evaluaciones de políticas públicas, hemos tenido casos exitosos y otros no tan exitosos. Entonces, hemos tenido casos en los que algunas recomendaciones de las evaluaciones han servido para que se tomen decisiones. Y también hay otras en las que nos pasa que quedan como a medias, (…) no sé si sería dividido por mitades, pero yo creo que sí está repartido, como repartido el caso en el que sí, el caso en el que no." (Entrevista contratista DSEPP1)

Las expresiones "ciertas evaluaciones que se han utilizado" o "algunas de las evaluaciones han servido", resulta significativo, no por el hecho de que brinde una medición exacta acerca de la efectividad de la política, o porque brinde una percepción en toda su complejidad (teniendo en cuenta los diversos usos que pueden tener las evaluaciones, además del que está reglamentado como objetivo

en los decretos). Resulta significativo porque esta valoración tímida acerca del uso de las evaluaciones está siendo realizada por quienes defienden abiertamente la política pública analizada.

Cabe aclarar que las percepciones que tienen los evaluadores/consultores acerca de la implementación de las recomendaciones que elaboran, o en general del uso diversificado que se le pueda dar a las evaluaciones que realizan, tienen un alcance limitado. La terminación del contrato antes del seguimiento a la implementación de las recomendaciones, y, por ende, la disolución en la coordinación entre los procesos, así como la escasa participación en los espacios de difusión de la evaluación, implican que estos desconozcan el detalle de lo que sucede con las evaluaciones que realizan.

2.3.2 Percepción general acerca de factores que apoyan o restringen el uso de las evaluaciones.

Por lo general, los cuestionamientos acerca de la utilización de las evaluaciones permitieron al mismo tiempo identificar la percepción acerca de los factores que restringen o que apoyan el uso de las evaluaciones. En cuanto a los factores que apoyan el uso se identificaron:

1. Calidad técnica de las evaluaciones (coordinadora DSEPP)
2. Continuidad de los funcionarios de las entidades sectoriales en los dos comités (coordinadora DSEPP; contratista DSEPP1)
3. Disposición para el trabajo conjunto de los funcionarios de las entidades sectoriales en los dos comités (contratista DSEPP1).
4. Apropiación de la evaluación por parte de las entidades sectoriales (coordinadora DSEPP).
5. Supervisión técnica del delegado del Comité de Seguimiento – secretaría técnica- (contratista DSEPP2)
6. Diversificación y selección acertada del alcance de la evaluación (contratista DSEPP1; consultor Econometría).

Como puede observarse, todos estos factores ya han sido identificados en las teorizaciones presentadas en el capítulo anterior, así que lo único que hacen es confirmar la importancia de estas variables. Por su parte, dentro de los elementos que subrayan que inciden de manera negativa en la utilización de las evaluaciones se encuentran:

1. Desconocimiento a largo plazo del estado de la implementación de las recomendaciones –ausencia de seguimiento sistemático- (coordinadora DSEPP).

2. Exceso de tecnicismo -perjudica lenguaje común entre sectores y, por ende, apropiación de la evaluación[46]- (Consultora del Centro Nacional de Consultoría).
3. Inflexibilidad del diseño de la evaluación -aumenta irrelevancia de los resultados- (Consultor del Centro Nacional de Consultoría).
4. Usuarios primarios indeterminados[47] (Consultor del Centro Nacional de Consultoría).
5. Ausencia de interés en la evaluación por parte de la entidad sectorial (contratista DSEPP 2; consultor Econometría[48]).
6. Limitada difusión de los resultados de la evaluación (consultora del Centro Nacional de Consultoría).
7. Usuarios primarios sin capacidad de decisión[49] (coordinadora del Grupo de Evaluaciones Estratégicas de la DSEPP).
8. Recomendaciones difíciles de implementar (coordinadora del Grupo de Evaluaciones Estratégicas de la DSEPP).
9. Inflexibilidad normativa –cuando las recomendaciones implican modificaciones normativas- (contratista DSEPP1; contratista DSEPP 2).
10. Capacidad institucional de entidades sectoriales insuficiente –carencia de personal y presupuesto- (contratista DSEPP1).
11. Actitud reticente hacia la evaluación –ausencia de cultura evaluativa- (contratista DSEPP 1; consultor Universidad de los Andes).
12. Hallazgos de la evaluación sustentan decisiones ya tomadas (consultora Centro Nacional de Consultoría; consultor Econometría).

Los primeros diez de estos factores que los participantes de las situaciones de acción de las evaluaciones reconocen como explicaciones de la no utilización de la

[46] Al referirse a este problema se manifiesta puntualmente: "… como viene la solicitud de términos de referencia de las evaluaciones de DNP en el diseño de las mismas, por esa exigencia de calidad técnica es que a veces se genera un exceso de tecnicismo. Esto hace que las evaluaciones no sean aterrizadas al lenguaje de los funcionarios de estas entidades y finalmente por eso se hace más difícil la apropiación también (…) no es por el tipo de recomendaciones que hacemos que pueda haber apatía, tampoco por falta de interés. Yo creo que es más por lo que dije antes, por un problema de lenguaje, de arrogancia técnica del DNP que hace que las evaluaciones y sus recomendaciones tengan un lenguaje lejano y que ellos [los funcionarios] no entiendan cómo pueden implementarlas" (Entrevista consultora de Centro Nacional de consultoría).

[47] Esta problemática se abstrae de: "… a veces pasa, como en estos ejemplos que mencioné, que no es tan claro quién es la contraparte en la entidad que pide la información o que se va a beneficiar de ella" (Entrevista consultor del Centro Nacional de Consultoría).

[48] Al respecto, este consultor sostiene: "[dificultades] hay varias. Una, es si la evaluación realmente se requiere o no. Nosotros somos consultores, y si nos contratan para contar cuántos pelos tiene esa alfombra, pues denos los recursos y listo. O sea, quisiéramos que todo lo que hacemos tenga incidencia (…) [pero] no necesitaban la evaluación" (Entrevista consultor Econometría).

[49] Puntualmente se comenta: "… al final las decisiones se toman en niveles directivos, niveles mucho más arriba de lo que… a niveles políticos" (Entrevista coordinadora del Grupo de Evaluaciones Estratégicas de la DSEPP).

evaluación, encuentran soporte empírico y teórico dentro del análisis institucional y de coordinación. El siguiente es un problema de cultura de la evaluación, la DSEPP aborda este problema a partir de recalcar la importancia de las evaluaciones en las distintas reuniones que se realizan con entidades sectoriales, y en todo caso, es un problema que debe ser abordado desde un análisis de los modelos mentales de los usuarios primarios con respecto a la evaluación.

El último problema está relacionado con la tendencia hacia el gerencialismo identificada por Guba y Lincoln. Las siguientes citas ayudan a caracterizarlo:

> “Entrevistada: No creo que las evaluaciones se usen para rendición de cuentas ni para mejora de los programas. En el caso de FEST [Programa Familias en su Tierra] si bien se tomó la decisión de acabar con el programa a partir de la evaluación, la decisión se tomó en el DNP no en Prosperidad Social.
> Entrevistador: O sea se utilizan un poco como para justificar las decisiones que ya han sido tomadas de antemano.
> Entrevistada: Sí, en muchas.” (Entrevista consultora del Centro Nacional de Consultoría).
>
> “Y llegamos a la conclusión de que los puntos digitales de la fase I (…) había que cambiarlos. Y ya iban en la fase III y ya los habían cambiado. No necesitaban la evaluación porque no iban a tomar decisiones (…) no servía para tomar decisiones porque esas decisiones ya se habían tomado dos años atrás” (Entrevista consultor Econometría).
>
> “… les sirvió para confirmar que sus decisiones habían sido acertadas al cambiar el tema de los puntos, como para darle tranquilidad a los funcionarios. Y si algún día los investigan, pues tienen la evaluación para demostrar que eso había que cambiarlo” (Entrevista consultor Econometría).
>
> “… en cuanto a Ser Pilo Paga (…) pienso que la evaluación no ha generado algún cambio sustancial en el enfoque del programa porque finalmente es un programa mucho más político, es decir, la decisión está basada puramente en una decisión política más que en cualquier otra cosa. Así pues, decir: "vamos a quitar este modelo para volverlo de otra manera", no creo que haya mucho la voluntad de hacerlo.” (Entrevista consultor del Centro Nacional de Consultoría).

Por un lado, puede pensarse que la evaluación, en efecto, está siendo utilizada al sustentar o documentar las decisiones que ya han sido tomadas, lo cual significaría la efectividad del proceso evaluativo. Sin embargo, el objetivo de la política pública de evaluación de políticas en Colombia (o en la literatura académica), no es brindar evidencia para fortalecer las decisiones ya tomadas de las entidades sectoriales, sino para racionalizar (o servir de insumo a) la toma de decisiones (Decreto 1082 de 2015, art. 2.2.7.1.1.1.). Luego de que se toma una decisión, esta se convierte en

una acción a implementar y no hay ocasión para decidir por algo mejor antes de implementar esa acción.

Una evaluación que sirve a modo de argumentación o justificación de decisiones ya tomadas (acciones a implementar o que ya se están implementando), no es una evaluación efectiva precisamente porque no está racionalizando o mejorando ninguna toma de decisión. En otras palabras, no es efectiva porque no informó ni sirvió de insumo al momento en el que se decide una particular elección, la cual es resultado y se identificó gracias a la nueva información de la evaluación. A pesar de lo intrincado que pueda parecer este asunto, su relevancia radica en que se refiere esencialmente a la diferencia entre racionalizar la gestión pública o buscar su legitimación, razón por la cual merece una particular atención.

2.3.3 Implementación de las recomendaciones de las evaluaciones a SINERGIA en los años 2010 y 2013

A continuación, se presenta una tabla que busca presentar la medición acerca del estado de implementación de las recomendaciones que se le realizaron al componente de evaluaciones estratégicas en las evaluaciones realizadas a SINERGIA en los años 2010 y 2013. Tal como se puede observar, el uso de la evaluación es un tema central que se aborda y se recomienda en ambas evaluaciones. Las respuestas que se ponen en la segunda columna están basadas en el análisis institucional realizado, en las entrevistas, y en las mediciones que se han realizado para el análisis comparado a julio de 2019.

Recomendación	Estado de imp.
Línea de base del acceso, uso y la calidad de la información que produce SINERGIA (Centro Nacional de Consultoría, 2010a).	
1. “Se solicita mayor difusión y publicidad a las evaluaciones. Puede ser mediante correos electrónicos”.	Implementada
2. “Incorporar otras áreas en las evaluaciones como infraestructura, competitividad, agricultura, seguridad y políticas en el sector minero y energético”.	Implementada
3. “Presentar las metodologías con mayor claridad, usando un lenguaje entendible por todo público”.	Implementada
4. “Publicación de todas las bases de datos y manuales de las mismas”.	No implementada
5. “Realizar la socialización de resultados con mayor periodicidad”.	Implementada

6. “Lograr formar un ciclo de realimentación positivo donde los programas son evaluados y las recomendaciones que surgen de la evaluación son llevadas a cabo”.	No implementada
Evaluación intermedia de la operación y los resultados del Sistema Nacional de Evaluación de Gestión y Resultados –SINERGIA- (Gerencial Ltda., 2013a).	
1. “Adoptar una agenda de mayor espacio temporal para las evaluaciones que permita hacer la selección de los temas a la luz de los lineamientos estratégicos del Plan Nacional de Desarrollo”.	No implementada
2. “Armonización de SINERGIA Seguimiento con SINERGIA Evaluaciones. Las evaluaciones priorizadas en la agenda deberían obedecer a las prioridades del Plan de Desarrollo”.	No implementada
3. “Analizar el historial de solicitudes de extensiones de plazos en los contratos y de ser necesario, introducir cambios en las directrices de definición de los tiempos asignados al trabajo de campo y al procesamiento”.	Implementada
4. “Metodologías de evaluación alternativas a las existentes”.	Implementada
5. “Plan de divulgación de evaluaciones para que lleguen a sectores más amplios de la población”.	No implementada
6. “Claridad en lenguaje de evaluaciones para mejorar difusión”.	Implementada
7. “Articular resultados de las evaluaciones con decisiones sobre presupuesto de inversión”.	No implementada
8. “Promocionar el uso efectivo de los resultados las evaluaciones en la toma de decisiones en dos niveles:” • “en el nivel de las asignaciones presupuestales en cabeza de la Dirección de Inversiones y Finanzas Públicas (DIFP) del DNP;” • “ii) en el nivel de la institución encargada del programa evaluado”.	No implementada
9. “Las evaluaciones serán efectivas si las recomendaciones se aplican. En este sentido la aplicación del documento “Plan de Transferencia e Implementación de Recomendaciones” debe ser aplicado y monitoreado”.	No implementada
10. “Llevar historial de evaluaciones. Revisar y actualizar la memoria de Evaluaciones y asignar responsables para actualizar el historial de las evaluaciones efectuadas y la verificación de efectos”.	Implementada
11. “Enfatizar el seguimiento a la aplicación de las recomendaciones, por lo menos en dos alcances: El acompañamiento a la aplicación de las recomendaciones derivadas de las evaluaciones contratadas y la verificación de los efectos derivados de dichas aplicaciones”.	No implementada
12. “Implementar un sistema informático de evaluaciones. Definir los requerimientos para un nuevo sistema informático que permita la gestión del proceso de evaluaciones”.	No implementada

Recomendaciones extraídas de las fichas técnicas (Centro Nacional de Consultoría, 2010b; Gerencial Ltda., 2013b) de estos procesos de evaluación, los cuales se encuentran en: https://sinergiapp.dnp.gov.co/#Evaluaciones/EvalFin/145

Ocho de dieciocho recomendaciones al componente de evaluaciones de SINERGIA han sido implementadas a 2019. Tal como se observa, en ambas evaluaciones se sugiere a la DSEPP profundizar su trabajo hacia el uso de las evaluaciones por parte de las entidades evaluadas, estableciendo en consonancia con esta investigación la equivalencia entre la implementación de las recomendaciones y la efectividad de los procesos de evaluación.

2.4 Conclusiones

El análisis institucional permite sacar conclusiones para cada tipo de reglas. En cuanto a las reglas de alcance, la normatividad jurídica y técnica direcciona satisfactoriamente el estándar de oportunidad, pertinencia y utilización. Con el estándar de respuesta sistemática a las recomendaciones y seguimiento de estas existe ambigüedad: por un lado, existen reglas que direccionan el estándar y direccionan su relevancia, pero por otro estas han sufrido internamente un proceso de transformación radical (remplazo de la Matriz de Gestión de Recomendaciones por la Matriz de Uso), que implicarían su no cumplimiento. Esta situación ambigua no descarta, en todo caso, que existan procesos de evaluación en los que se cumpla este estándar dependiendo de la importancia de la política para el DNP.

El estándar de desarrollo de capacidad no está direccionado por ninguna normatividad. En términos de Patton, este estándar correspondería a un tipo particular de uso de proceso (2012, p. 140), los cuales corresponden a las cosas que las personas involucradas en una evaluación aprenden por estar en ese proceso de evaluación, como usar la lógica de la evaluación, emplear procesos de razonamiento sistemático al hacer análisis, o ser guiados por los estándares que informan la práctica evaluativa. Su diferencia con respecto al uso de los hallazgos es equivalente a aprender a cómo aprender frente a aprender conocimiento sustantivo sobre algo. Debido a la complejidad de estos usos de proceso, y del estándar de desarrollo de capacidad, la inexistencia de normatividad al respecto no permite descartar que pueda encontrarse esto en los contratos de evaluación.

De los tipos de reglas que se pueden encontrar en la normatividad general a todos los proceso de evaluación, las que corresponden a las reglas de posición y a las de frontera, son las que mejor abordan los estándares profesionales de calidad de la OCDE, puntualmente se puede decir que existen reglas y normas claras que direccionan los estándares de participación de los actores relevantes; consideración

de evaluación conjuntas; enfoque asociativo; estructuras de gobierno y de gestión; e independencia del evaluador[50].

A pesar de lo anterior, es de resaltar que la reglamentación general de la política brinda al DNP diversas competencias que en la literatura académica se asignan a los evaluadores. Algunas de estas competencias son: difusión de la evaluación, identificación e involucramiento de los usuarios previstos, asesoramiento metodológico a los usuarios previstos, apoyo para la definición del alcance de la evaluación y las preguntas evaluativas y seguimiento a la respuesta e implementación de las recomendaciones.

En general se observa que, en toda la etapa de diseño de la evaluación, es la DSEPP la que tiene el rol de evaluador. Esto sucede al mismo tiempo que puede tener roles de intermediario en la comunicación entre evaluador y usuarios primarios, supervisor de la evaluación, e incluso solicitante (cliente) de la evaluación. Esto puede significar falta de claridad en los roles y en la distribución de responsabilidades, y más relevante aún, en la definición de los usuarios previstos de la evaluación (estándar de participación de actores relevantes), y, por ende, de los usos que se le van a dar a esta (su propósito).

El tipo de reglas que más estándares incluye en su interior corresponde a las reglas de elección. Con respecto a los estándares que determinan las situaciones de acción correspondientes al diseño de las evaluaciones, el marco normativo de la política direcciona satisfactoriamente los estándares de evaluabilidad; coordinación y alineamiento; objeto y alcance de la evaluación; objetivos específicos de la evaluación; cuestiones a responder; de selección y aplicación de criterios de evaluación; selección de enfoque y metodología; realización de la evaluación dentro del plazo y el presupuesto asignados; y fundamentación y finalidad de la evaluación.

Se debe hacer la salvedad que en cuanto a este último estándar, no se encuentran requerimientos puntuales acerca del establecimiento de la utilización prevista, y a pesar de poder deducirse de la determinación de las preguntas y objetivos a partir de una demanda de información cubre este punto, y de la solución de los llamados “cuellos de botella”, no hay una definición puntual de la utilización prevista, lo cual, también está en correspondencia con el análisis del trabajo conjunto sobre los usos de la evaluación.

Con respecto a los estándares que componen reglas de elección para las situaciones de acción de ejecución de la evaluación, primero hay que establecer que el estándar “informe de la evaluación” se descompone en un conjunto particular de estándares de calidad, por lo que su medición debe hacerse a partir de la

[50] El estándar de independencia de los evaluadores es direccionado por la normatividad jurídica y técnica, sin embargo, esto no implica que en efecto el evaluador sea independiente, tal como lo revela la discusión de reglas de agregación y de coordinación.

agregación de estos últimos. Estos estándares son: explicación de la metodología; lógica de la intervención; validez y fiabilidad de fuentes de información; mención de discrepancias entre miembros del equipo evaluador; mención de los cambios y limitaciones de la evaluación; y contexto de la intervención; respuesta a las preguntas evaluativas; claridad del análisis; y claridad y representatividad del resumen.

De estos nueve estándares únicamente los últimos dos cuentan con reglamentación transversal a todos los procesos de evaluación. Los demás estándares son institucionalizados en los documentos de estudios previos, en el pliego de condiciones definitivo para la contratación, y en el contrato de consultoría de cada contrato de evaluación en particular, en este sentido, la ausencia de reglamentación general para estos estándares no significa que deban ser descartados de la presente investigación. Lo mismo sucede con el estándar de consulta y protección de los implicados, el cual también se reglamenta en cada contrato en particular, pero no se desprende conceptualmente del estándar de informe de evaluación.

Los estándares que corresponden a las reglas de agregación están todos direccionados a partir del marco institucional general de la política, sin embargo, esta institucionalidad deja abierta la posibilidad a que exista una tensión entre estos estándares de calidad. La reglamentación acerca de los estándares de control de calidad y de inclusión de observaciones de los implicados entran en conflicto con el estándar de proceso evaluativo libre y abierto, esto es así porque la fuerte injerencia que pueden tener las entidades sectoriales y el DNP en los productos (como el hecho de que se requiera su aprobación para el pago al evaluador) si bien, no elimina la discrecionalidad del evaluador, si facilita lo que en el análisis de la percepción de la efectividad de la política se denominó como gerencialismo.

Al no existir estándares que correspondan a reglas de pago, el último tipo de reglas restante son la de información. Todos los componentes del estándar de documentación de finalidad y expectativas se incluyen en la reglamentación general. En cuanto al estándar de difusión, a pesar de que dentro de la reglamentación general no incluya la participación del evaluador contratista en este aspecto, todos los componentes del estándar de difusión (formato accesible, difusión sistemática, difusión externa e interna) se encuentran incluidos en esta reglamentación.

En síntesis, los estándares de calidad que presentan debilidades son: respuesta sistemática a las recomendaciones y seguimiento; desarrollo de capacidad; participación de actores relevantes; fundamentación y finalidad de la evaluación; y el de proceso evaluativo libre y abierto. A pesar de estas debilidades, la reglamentación general no va en contravía de estos, ni los descarta de manera tajante, por tal motivo, no puede descartarse la relevancia de su estudio o que en los casos en los que sea positiva la medición de efectividad del proceso de evaluación estos factores no existan.

El estándar de ética de la evaluación es un caso especial, por un lado, correspondería a una regla de posición en tanto que determina el rol general del evaluador, sin embargo, al ser tan abstracto en su contenido, y al no existir reglas que lo direccionen, se puede concluir que lo mejor que se puede hacer con este estándar es abordarlo como componente de la estructura general de la comunidad (variable explicativa de la arena de acción en el IAD de Ostrom), por lo que no es procedente abordarlo como algún tipo de regla puntual dentro de la situación de acción. En este sentido, sería el único estándar a descartar para el abordaje del análisis comparado.

El análisis general de la coordinación, en su conjunto, permitió evidenciar muchas más fallas en el proceso de evaluación que el análisis general institucional. En cuanto al involucramiento de los usuarios previstos, si bien se ha diseñado todo el proceso de evaluaciones estratégicas con un particular cuidado de analizar e incluir a los participantes relevantes, las entrevistas evidencian que, a pesar de esta institucionalización, la coordinación se puede ver restringida por la falta de claridad acerca de quiénes los usuarios primarios. También se puede ver restringida en casos en que no se convocan en los comités de diseño o seguimiento a todos los actores de la política pública, razón por la que se genera un déficit de capacidad en los usuarios primarios para el uso de la evaluación[51].

Una vez identificados los usuarios previstos, estos y los evaluadores deben conformar un equipo de trabajo conjunto en el cual se desenvuelvan interacciones de naturaleza activa-reactiva-interactiva-adaptativa tal como se expuso en el modelo analítico. En este punto, se refleja que el elevado grado de control que tiene la DSEPP en la coordinación que existe entre evaluador y usuarios previstos, incide de manera negativa en la fluidez de los ajustes mutuos que se deben dar entre estos, y en los ajustes de naturaleza reactiva[52] que debería realizar la DSEPP para responder a la demanda de los usuarios previstos. Esto último se manifiesta en las barreras en el lenguaje y la comunicación por un elevado grado de tecnicismo[53] que han sido identificadas por los evaluadores/consultores.

---

[51] En un escenario en el cual cada vez las políticas públicas son más intersectoriales, y necesitan cada vez más espacios de coordinación, la convocatoria incompleta de usuarios primarios genera fallas en el proceso de coordinación, como fallas en la identificación conjunta de preguntas relevantes, en la interpretación de hallazgos, o en el estudio de factibilidad de las recomendaciones, y probablemente en muchos otros componentes de la coordinación.

[52] Es importante recordar que para Patton se es reactivo en escuchar a los usuarios previstos y en responder a lo que ellos aprendan acerca de la particular situación en la que se desenvuelva la evaluación (2012, p. 47).

[53] Este problema en la interacción comunicativa entre los sectores de gobierno ya ha sido caracterizado por Habermas en su obra Facticidad y Validez, en esta obra lo aborda como un problema de integración social en donde los sistemas funcionales desarrollan su autonomía a la par que desarrollan sus propios códigos y sus propias semánticas, no traducibles las unas a las otras, con esto, los sistemas pierden su capacidad de comunicarse, se ven cada vez más lejos los unos de los otros y quedan a la vez más de espaldas unos a otros, surge entonces una improbabilidad e

Frente a la reacción del evaluador a las preocupaciones y demandas de usuarios, se debe partir del hecho de que no es el evaluador el que tiene la oportunidad de reaccionar y ajustarse a las demandas y preocupaciones de los usuarios previstos, sino que es la DSEPP en el Comité de Diseño quien caracteriza esta demanda, y lidera su abordaje a partir de una metodología de cadena de valor y de identificación de "cuellos de botella". Esto significa que el ajuste que realiza el evaluador/consultor a la demanda es únicamente de adaptación a las decisiones ya tomadas, por lo que no puede asesorar o apoyar la caracterización de esta demanda, o manifestarse de manera oportuna acerca de la factibilidad de la responder a esa demanda.

Esta escasa participación del evaluador/consultor también aborda el trabajo conjunto sobre los objetivos y usos previstos, el cual se da únicamente con coordinación entre la DSEPP con las entidades sectoriales. No se encuentra ni reglamentación, ni evidencia en las entrevistas acerca de discusiones puntuales acerca de los usos de la evaluación, aun así, el uso implícito que existe es la solución a los cuellos de botella en la cadena de valor reconstruida. Con los debates metodológicos pasa algo similar, la participación del evaluador/consultor es muy restringida, especialmente en las etapas tempranas de diseño de la evaluación, en donde la DSEPP busca cubrir esta ausencia del evaluador, siendo ellos quienes abordan funciones características del evaluador como el asesoramiento metodológico.

Desde el U-FE, estos debates metodológicos deben abordarse de manera continua con el fin de adaptar conjuntamente el diseño de la evaluación a las transformaciones en la demanda de los usuarios primarios. Se evidencia que el diseño de la evaluación tiene una naturaleza rígida a pesar de que las evaluaciones pueden durar hasta dos años y medio desde que se caracteriza la demanda hasta que se tienen resultados. Esto genera fallos en la respuesta a la demanda de los usuarios y en la apropiación que estos realizan a la evaluación, principalmente por la pérdida de relevancia que aumenta con el tiempo.

En cuanto al trabajo conjunto en la interpretación de hallazgos y construcción de recomendaciones, a pesar de las consideraciones que se han realizado acerca del elevado control de la DSEPP en los escenarios de coordinación, así como de los riesgos que esto implica para la independencia del evaluador, se puede constatar que tanto el análisis institucional como las entrevistas revelan que existe un diálogo y un ajuste mutuo para la interpretación de los hallazgos y para la construcción de las recomendaciones.

---

inverosimilitud del logro de comunicación entre unidades autónomas que operan autorreferencialmente (1998. pp. 415).

En la planeación conjunta del seguimiento a la implementación de las recomendaciones, la participación de los evaluadores termina con la formulación y corrección del Plan de Transferencia e Implementación de Recomendaciones. Este plan brinda los temas a los cuales se les debe realizar el seguimiento, a pesar de que no constituye un plan de seguimiento como tal, puede reflejar elementos en esta dirección.

La DSEPP manifiesta la realización de este seguimiento, en estos momentos es una actividad unilateral que no tiene en cuenta a los usuarios previstos (tal como se observa con la utilización de la denominada “Matriz de Uso”). También se evidencia un desconocimiento intencional de la responsabilidad de la DSEPP con respecto a la incidencia en las decisiones de las entidades sectoriales, el cual se utiliza a modo de justificación para la no realización de un seguimiento sistemático a la implementación de las recomendaciones.

No existe una coordinación a partir de la búsqueda proactiva del uso por parte del evaluador, ya que su papel culmina con la presentación de los productos que requiere su contrato de consultoría y su participación en la difusión de la evaluación es muy restringido. La DSEPP, a pesar de asumir en múltiples momentos funciones que debería ejercer el evaluador, al terminar la evaluación no se apropia de estas con el objeto de ejercer una defensa activa de los resultados de las evaluaciones, en parte porque consideran que no es su competencia incidir en las decisiones de las demás entidades (tal como se revela con datos de entrevistas), a pesar de que los decretos reglamentan lo contrario. Aun así, la información recolectada hasta el momento no permite descartar que en algunos casos el DNP o los contratistas realicen una búsqueda proactiva del uso.

Finalmente, frente a la presentación de la percepción general de los resultados de la Política Pública de Evaluación de Políticas, se destaca una valoración tímida por parte de la misma DSEPP acerca del uso de las evaluaciones y del impacto que la política ha tenido en las decisiones de las entidades sectoriales. También que todos los factores identificados como promotores del uso se encuentren en completa correspondencia con alguna de las dos teorizaciones acerca de la efectividad de los procesos de evaluación seleccionadas.

Por último, los funcionarios de la DSEPP y los evaluadores/consultores identificaron doce factores como restrictivos del uso de las evaluaciones. Diez de estos encuentran soporte empírico adicional en el análisis institucional y de coordinación realizados previamente en este mismo capítulo. Los últimos dos se refieren a un problema de naturaleza cultural, y a que la evaluación se utilice a modo de argumentación o justificación de decisiones ya tomadas, con lo cual se perdería la finalidad de las evaluaciones de política pública, y se verían reducidas a un ejercicio de búsqueda de legitimidad a partir de la manipulación de un ejercicio técnico. Estos últimos dos problemas deben tomarse en cuenta como variable explicativa de la situación de acción “estructura de la comunidad”; y como una falencia en la coordinación para la construcción conjunta de las recomendaciones, respectivamente.

Finalmente, se realizó una pequeña medición acerca del grado de efectividad que tuvieron las evaluaciones realizadas al componente de evaluaciones de SINERGIA en 2010 y 2013, en el cual se pudo establecer que solo ocho de dieciocho recomendaciones han sido implementadas. La medición realizada a la efectividad de los demás procesos de evaluación se realizará de manera similar, sin embargo, se profundizará mucho más en su argumentación, lo cual no se consideró necesario en este capítulo, teniendo en cuenta que las mediciones de todos los capítulos de casos ayudan a caracterizar, en su conjunto, el grado de implementación de estas recomendaciones.

El análisis realizado en este capítulo no tiene el alcance ni el apalancamiento analítico necesario para poder comprobar o falsear las teorizaciones, sin embargo, permite realizar un avance exploratorio significativo al determinar la correspondencia que existe entre el diseño general de la política y los estándares profesionales de calidad, así como el grado de complejidad de la coordinación de los actores. Esto brinda algunas ideas preliminares acerca de la naturaleza de los problemas que explican la variabilidad en la efectividad de la política pública, y permite focalizar su medición en los siguientes capítulos. Por último, este capítulo ha servido a su propósito principal, permitiendo aclarar que las variables seleccionadas pueden ser estudiadas en el caso colombiano, y son relevantes para analizar la efectividad de los procesos de evaluación.

3. Efectividad de la evaluación de impacto y de diseño al Programa de Conexiones Intradomiciliarias del Ministerio de Vivienda, Ciudad y Territorio

La estructura de los capítulos en los que se presentaran los casos a comparar, y, por lo tanto, de este, es la siguiente: en primer lugar, se va a realizar una breve presentación de la política pública, esto con el propósito de contextualizar el alcance, los objetivos y los resultados de la evaluación. Posteriormente, se realiza la medición de cada uno de los estándares de calidad organizados por cada uno de los componentes de la situación de acción y las reglas que los determinan. Se continúa con un análisis de los patrones de interacción a partir de los componentes de la coordinación como ajuste mutuo. Por último, se realiza una medición de la efectividad del proceso de evaluación en donde se mide el grado de implementación de cada una de las recomendaciones de la evaluación, y si estás buscan o no justificar decisiones tomadas previamente a la evaluación.

3.1 Presentación del Programa de Conexiones Intradomiciliarias.

El objetivo principal del Programa de Conexiones Intradomiciliarias (PCI) es “fomentar el acceso a los servicios públicos domiciliarios de acueducto y alcantarillado, mediante la construcción o el mejoramiento de las conexiones intradomiciliarias” (Centro Nacional de Consultoría –CNC-, 2015a, p. 3). Este objetivo ha sido enmarcado por parte de FONADE (Fondo Financiero de Proyectos de Desarrollo) en el mejoramiento del Índice de Pobreza Multidimensional y en el alcance de los objetivos de la RED UNIDOS. Particularmente, observando los objetivos de la RED UNIDOS, el subsidio del PCI para hogares de estratos 1 y 2, se enmarca en 3 de sus 45 logros básicos (FONADE, Min. Vivienda, 2012, p. 2):

- Logro 21: “La Familia practica hábitos saludables en la manipulación y preparación de alimentos”.
- Logro 24: “La vivienda cuenta con abastecimiento de agua potable y un sistema de desagües”.

- Logro 27: “La familia cuenta con espacios diferenciados de baño, cocina, lavadero y dormitorios”.

El PCI también se enmarcó en el Plan Nacional de Desarrollo de 2010-2014 “Prosperidad para Todos”, en cuanto que este buscaba “Consolidar las reformas estructurales en el sector de agua potable y saneamiento básico, para lograr impacto positivo en la disminución de pobreza a través de coberturas reales de acueducto, alcantarillado y aseo” (DNP, 2015a, 152, p. 2), el cual también tenía por objetivo específico en esta materia, la construcción de conexiones intradomiciliarias para la población más pobre, considerando una meta de 90.000 conexiones intradomiciliarias construidas para 2014.

Dentro del diseño de esta intervención se considera en cuanto a la conexión intradomiciliaria: el suministro e instalación de hasta 5 equipos y aparatos de bajo consumo, los cuales incluyen inodoro, lavamanos, lavaplatos, lavadero y ducha, es incluye la construcción de redes y accesorios. Considera la instalación de un tanque de 500 litros cuando sea necesario (en donde la presión del agua es baja o donde el servicio es intermitente), una cajilla de inspección y adecuaciones necesarias para la funcionalidad. Considera los acabados, el retiro de escombros y pruebas de presión y flujo (MVCT, 2013, p. 6). La conexión domiciliaria sólo se realiza cuando se evidencia necesidad de intervención intradomiciliaria, y esta consta de la instalación de acometidas de acueducto y alcantarillado, y de la instalación del medidor de acueducto.

El PCI es un programa que busca incidir directamente en las condiciones de agua, saneamiento e higiene (ASH) de las viviendas, con lo cual se espera que exista una menor incidencia de enfermedades asociadas con el consumo de agua contaminada (Enfermedades Diarreicas Agudas – EDA), también que exista y una menor reproducción de vectores y una adecuada eliminación de excretas (CNC, 2015, p. 40). Una de las hipótesis principales del PCI es que “proveer a los hogares con sanitarios, lavamanos, lavaplatos, lavaderos y duchas puede ayudar a romper con el ciclo de generación y transmisión de gérmenes y bacterias patógenas” (p. 5).

La lógica causal de la intervención parte del problema principal/central, que corresponde al alto índice de hogares con condiciones inadecuadas para el acceso a agua y saneamiento básico. Esto es causado por deficiencias al interior de las viviendas en relación con el acceso al agua potable y el manejo de aguas residuales, por las malas prácticas de higiene y uso del agua, y por la baja cobertura en servicios públicos de acueducto y alcantarillado.

La primera de estas causas, a su vez, es determinada por la predominancia de la utilización de materiales inadecuados para la construcción de conexiones intradomiciliarias, y por la ausencia de sistemas domiciliarios; la segunda es determinada por el desconocimiento de la población en temas de salubridad y por prácticas culturales inadecuadas de autocuidado; finalmente, la tercera es causada

por la insuficiencia en infraestructura para la prestación de estos servicios. El Ministerio de Vivienda, Ciudad y Territorio (MVCT) esperaba producir impactos cuantificables en variables como la salud física, salud psicosocial, derechos sexuales y reproductivos, crecimiento y desarrollo infantil, salud ambiental, déficit cualitativo de vivienda, productividad, asistencia escolar, uso racional del agua y pobreza (2015a, p. 3).

3.2 Caracterización de las situaciones de acción del proceso de evaluación del Programa de Conexiones Intradomiciliarias.

3.2.1 Reglas de alcance y resultados potenciales

Esta evaluación fue realizada entre diciembre de 2014 y septiembre de 2015. Los estándares de calidad que entran en esta categoría tienen la particularidad de que funcionan como reglas de elección en la situación de diseño de la evaluación, pero como reglas de alcance que establecen los resultados potenciales en la situación de ejecución de la evaluación. En primer lugar, con respecto al estándar de fundamentación y finalidad de la evaluación, en el Pliego de Condiciones Definitivo para la contratación de la evaluación, en los Estudios Previos y en la Resolución 3789 de 2014 (por la cual se ordena la apertura del proceso de selección por concurso de méritos la evaluación del PCI), se establece:

> "Una evaluación de impacto (…) permitiría identificar los beneficios asociados directamente a este tipo de intervención, lo cual arroja evidencia suficiente para soportar la inversión pública enfocada a este tipo de programas, señalando si realmente estas intervenciones repercuten no solamente en el mejoramiento de la salud pública, sino también en otras áreas de interés para el desarrollo social, tales como la productividad, la asistencia y el rendimiento escolar y, en general, aquellas vinculadas al mejoramiento de la calidad de vida (…) es evidente la necesidad de contratar la evaluación del Programa, ya que para el desarrollo económico y social del país representa una oportunidad muy importante conocer los avances del sector y realizar ajustes a la política que aseguren con mayor precisión el logro de los objetivos" (DSEPP-DNP, 2014c, p. 3).

Conocer los avances del sector y realizar ajustes a la política son usos sumativos y formativos correspondientemente. Es claro para qué se realiza la evaluación, para quién se realiza, la razón por la que se hace en ese punto particular del tiempo y el modo en el que se pensó utilizar posteriormente los resultados, por lo que todos los componentes del estándar de fundamentación y finalidad de la evaluación se cumplieron en esta evaluación.

En cuanto al objeto y alcance de la evaluación, se resalta que un aporte del proceso de evaluación fue la reconstrucción y aclaración del modelo teórico del PCI a partir de definir la cadena de valor de la política pública. Esta fue la principal herramienta

que ayudó al MVCT a determinar los indicadores necesarios para medir el impacto y el diseño del PCI. Para el caso del PCI la cadena de valor es la siguiente:

Diseño	Insumos	Procesos	Productos	Resultados	Impactos
-Priorización -Estructuración -Convocatoria -Licitación -Contratación	-Aparatos hidrosanitarios -Materiales de construcción -Profesionales de construcción -Trabajadores sociales -Oficinas -Bodegas -Vehículos -Bases de datos	-Socialización -Obras -Capacitacio-nes -Acompañami-ento -Interventoría	-Unidades Sanitarias -Conexiones -Tanques -Lavaderos -Lavaplatos -Sistema de información	-Procedimiento mejorado de eliminación de excretas -mejores hábitos higiene -Mejoras en la autoestima -Cumplimiento de derechos -Facilidades en el acceso al agua	-Salud -Productividad -Asistencia escolar -Crecimiento y desarrollo infantil -Formalización de Viviendas -Beneficios psicosociales -Pobreza -Derechos sexuales
DNP, MVCT, FONADE	FONADE, contratistas, proveedores, cogestores, empresas de servicios públicos, EE.TT.	Interventoría, contratistas, proveedores, EE.TT., ESP	Hogares, ESP, EE.TT.	Hogares	Hogares, ESP, EE.TT.

Fuente: CNC, 2015b, p. 130.

Entre las ventajas que evaluadores y actualmente MVCT reconocen en la utilización de la metodología de cadena de valor, está la que permite entender la secuencia lógica de la intervención; permite entender el cómo para lograr el qué; y facilita la definición de indicadores mediante la esquematización de objetivos en cada una de las fases (CNC, 2015b, p. 123). En este sentido, fue satisfactoria la definición de la intervención u objeto a evaluar.

Entre los elementos de esta cadena, se seleccionó como elementos a abordar por parte de la evaluación los componentes de diseño e impactos de la intervención, así se describe este alcance:

> "(i) los impactos atribuibles a la intervención, a partir de la comparación que se realice entre los resultados para el grupo de población tratada por el Programa y los del grupo de población no tratada por el Programa; y (ii) las brechas de diseño del Programa, analizando la consistencia entre su estructura y los resultados de la intervención" (DSEPP-DNP, 2014c, p. 53).

En el informe final de la evaluación se resalta que la evaluación de impacto permite determinar si el programa está funcionando o no (si produce efectos en el bienestar de las personas), lo que ayudaría a las decisiones entorno a la expansión, continuación o eliminación de la intervención (p. 8). La evaluación de diseño se caracteriza como "complemento" de la evaluación de impacto, la cual permite

identificar cuellos de botella y fallas en el diseño que impiden que el programa alcance los resultados esperados, por lo que permite "…convertir los resultados de la evaluación de impacto en recomendaciones concretas para mejorar la efectividad de un programa" (p. 13). Esto significa que todos los componentes del estándar de objeto y alcance de la evaluación fueron aplicados satisfactoriamente.

Los objetivos específicos se aclaran en los Estudios Previos y en el Pliego de Condiciones Definitivo (p. 53), por lo que el estándar correspondiente también fue satisfecho completamente. Los objetivos específicos de la evaluación fueron:

1. Caracterizar a la población tratada y no tratada por el Programa en las áreas de interés identificadas en los estudios previos y en otras que identifique la consultoría, de acuerdo con los grupos de observación definidos para la evaluación.
2. Identificar, en la medida de lo posible y por medio de preguntas retrospectivas y/o bases administrativas, los cambios ocurridos en los hogares beneficiados y no beneficiados por el Programa en las áreas de interés: salud (enfermedades trasmitidas por consumo y manejo de agua de fuente no mejorada; enfermedades trasmitidas por manejo inadecuado de excretas), derechos sexuales y reproductivos, calidad de la vivienda, asistencia escolar, productividad, salud ambiental, pobreza, uso racional del agua, entre otras que identifique la consultoría.
3. Proponer el diseño de la evaluación cuasi-experimental más adecuado para determinar los impactos atribuibles al Programa.
4. De acuerdo con el diseño propuesto, identificar el efecto causal que el Programa tiene sobre la población beneficiada.
5. Proponer estrategias cuantitativas para identificar los principales mecanismos o canales de transmisión a través de los cuales el Programa impacta a los beneficiarios.
6. Identificar las brechas de diseño de la intervención en relación con los productos que entrega y los resultados esperados.
7. Identificar las brechas del diseño actual de la intervención en relación con el flujo de relaciones inter e intra institucionales que debe considerar el Programa para potenciar los resultados esperados.
8. Formular recomendaciones basadas en la evidencia, sobre la pertinencia de asignar recursos a este tipo de intervenciones y la manera cómo debería expandirse el Programa a nivel nacional.
9. Formular recomendaciones sobre el diseño de la intervención, teniendo en cuenta los estándares de política establecidos por la Organización Mundial de la Salud (OMS) y el flujo de interacciones que debe considerar el diseño para optimizar los resultados esperados.
10. Formular recomendaciones para mejorar el seguimiento de indicadores del Programa y la evaluación de este.

Frente a los componentes del estándar de oportunidad, pertinencia y utilización que pueden considerarse como reglas de alcance (los otros se verán en difusión, en

reactividad/coordinación o en efectividad de la evaluación), en el Pliego de Condiciones Definitivo para la contratación del evaluador, en los Estudios Previos para la contratación, en la Resolución 3789 de 2014 de apertura del concurso de méritos, en la Resolución 4219 de 2014 de adjudicación del proceso de selección por concurso de méritos, y en el Contrato de consultoría entre el DNP y el Centro Nacional de Consultoría, se define (con exactamente las mismas palabras) la necesidad que justifica la contratación.

Con esta definición de la necesidad se busca: "… resaltar la relevancia del sector para el desarrollo del país y la conveniencia de iniciar prontamente con la contratación de una evaluación de esta intervención que muestre los impactos de esta política pública y se puedan tomar decisiones informadas para asignar recursos y mejorar el diseño e implementación de la misma" (DSEPP-DNP, 2014c, p. 2). Como se observa, tanto en la finalidad en los objetivos específicos y en esta aclaración de la necesidad, se direcciona la evaluación a satisfacer necesidades de información para la toma de decisiones por parte del MVCT.

Así mismo, las conclusiones y recomendaciones se presentan de manera clara y orientadas a su puesta en práctica al interior del Plan de Transferencia e Implementación de Recomendaciones (PTIR), estas serán presentadas en la sección correspondiente a la medición de la efectividad del proceso evaluativo. La evaluación fue realizada en el cronograma establecido en el contrato de consultoría, y sus productos son sistemáticamente difundidos gracias al portal de internet de SINERGIA.

Lo anterior sumado a la labor del Comité de Diseño de la evaluación, pueden generar la impresión de una satisfacción completa de todos los componentes del estándar de oportunidad, pertinencia y utilización, sin embargo, tal como se podrá observar en la sección en el que se presenta la coordinación del proceso de evaluación con el de política pública, existieron fallas en cuanto a la reactividad que el evaluador y el DNP debieron tener frente a las necesidades y demandas de los usuarios primarios. Por esta razón el estándar solo fue parcialmente satisfecho.

Por último, frente a las reglas de alcance y los resultados potenciales de esta evaluación, queda por revisar el estándar de desarrollo de capacidad referente a usos de proceso en términos de Patton. Si bien no existen objetivos claros que evidencien el direccionamiento puntual de este estándar, la reconstrucción de la lógica de la intervención y la determinación de sus indicadores evidencian que a pesar de que no fueron considerados el desarrollo de capacidad, el proceso de evaluación aumento la capacidad del MVCT de medir su propia gestión. El coordinador del PCI de MVCT manifestó al respecto:

> "Gracias al trabajo del DNP se vio que hay unos impactos adicionales que nosotros de alguna manera los identificamos en campo, pero no sabíamos cómo medirlos. Todo el tema psicológico, el autoestima, y sobre todo un tema

> de la dignidad de la mujer relacionado con el espacio separado de baño, eso lo vimos en campo (…) estamos concentrando mucho el tema de estos impactos desde el punto de vista psicológico y también el tema de enfermedades diarreicas agudas, que ese también es otro impacto importante que nosotros no lo teníamos medido y que el DNP lo midió, que identificó que hay una disminución en el riesgo de padecer enfermedades diarreicas agudas, con unos porcentajes dependiendo de la población, sobretodo en menores (…) como valor agregado de esa evaluación, fue esos aspectos que nosotros nos imaginamos que podían impactar pero que no teníamos cómo medirlos, y esos resultados nosotros los utilizamos mucho, porque este programa se está promocionando mucho, y el programa ha sido incluso elegido por el Banco Mundial como experiencia exitosa a nivel mundial" (Coordinador PCI, MVCT).

No puede sostenerse que el DNP o el evaluador buscaran "maximizar los efectos positivos del proceso de evaluación sobre la capacidad evaluativa" del MVCT tal como lo requiere el estándar de calidad. Aun así, únicamente siguiendo el proceso normal de diseño de la evaluación, el DNP aumentó la capacidad del MVCT de medir los resultados de su gestión, por esta razón, puede decirse que el estándar está parcialmente satisfecho.

3.2.2 Reglas de frontera y participantes

Como se mencionó en el capítulo anterior, los espacios más importantes a observar para poder caracterizar la participación de actores relevantes a lo largo de todo el proceso de evaluación son los comités de diseño y de seguimiento. En los procesos internos de diseño y ejecución de la evaluación se establece que en cada reunión de estos comités debe quedar un acta. Estas actas fueron solicitadas por medio de derecho de petición a la DSEPP, su respuesta incluye lo siguiente:

> "… no hay actas de reunión entre los comités técnicos de evaluación con los consultores /evaluadores, debido a que como resultado de las reuniones se diligencia el formato de control de asistencia FGBH-09, en el cual hay información personal de los consultores (…) por lo tanto, no es posible atender esta solicitud" (DNP, respuesta al derecho de petición N° 20183600182751).

En todo caso, MVCT como líder del diseño e implementación del PCI si participó tanto en el Comité de Diseño como en el Comité de Seguimiento tal como se evidencia en las entrevistas a dos de sus funcionarios y en la respuesta al derecho de petición del DNP N° 20193600428271. Algunos actores importantes de la política como los entes territoriales, las empresas de servicios públicos y los beneficiarios solo tuvieron una participación de naturaleza consultiva al interior de este proceso, mientras que otros encargados de implementar recomendaciones como el Ministerio de Salud o el SENA no fueron incluidos en ninguna etapa de la

evaluación. Esto significa que los estándares de participación de los actores relevantes, enfoque asociativo y consideración de evaluación conjunta no fueron satisfechos en este proceso evaluativo.

Por parte de los evaluadores, es el consultor quien tiene la obligación de disponer del equipo de trabajo que es especificado en el Pliego de Condiciones Definitivo (DSEPP-DNP, 2014c, p. 4). Este equipo debe contar mínimo (pp. 59-60) con un director del proyecto que sea experto en evaluación de impacto (requerida experiencia de 2 años en asesoría o coordinación de estudios, consultoría o evaluaciones de impacto de políticas públicas en las que se utilizaran modelos econométricos); un experto en desarrollo humano y pobreza (con 2 años en asesoría o coordinación de estudios, consultorías o evaluaciones de políticas relacionadas con medición de indicadores de desarrollo humano y calidad de vida); un analista cuantitativo (dos años de experiencia en uso de técnicas cuantitativas para análisis de información estadística y uso de modelos econométricos); y un analista cualitativo (dos años en uso de técnicas cualitativas para la recolección de información y métodos cualitativos para el análisis de información).

El equipo evaluador fue contratado a partir de un proceso de selección abierto y transparente, con criterios habilitantes claros de naturaleza jurídica, financiera y técnica (dentro de estos últimos estaban el enfoque metodológico, el recurso humano, experiencia del interesado y apoyo a la industria nacional). En este proceso se recibieron siete propuestas, de las cuales cuatro estuvieron habilitadas para ser contratadas y, el Centro Nacional de Consultoría recibió la mejor calificación por enfoque metodológico. Con todo esto, el estándar correspondiente a equipo de la evaluación fue satisfecho en todos sus componentes.

El personal y las habilidades técnicas estuvieron satisfechas. Así mismo, en los requisitos financieros habilitantes se encuentra el índice de liquidez, índice de endeudamiento e índice de cobertura de intereses de la firma consultora, los cuales buscan medir la capacidad financiera de los proponentes, estimando así su capacidad de asumir compromisos y de cumplir oportunamente con las obligaciones del contrato (p. 19). El valor del contrato fue de $967.778.000 para el DNP. Esto significó que el estándar de recursos quedó completamente satisfecho.

3.2.3 Reglas de elección y acciones

Tal como se mencionó anteriormente, la intervención fue adecuadamente definida partir de una metodología de cadena de valor en donde participaron los usuarios previstos de la evaluación. El análisis evaluativo del diseño de esta cadena de valor es presentado en todo el capítulo correspondiente a la evaluación de diseño del PCI. El encadenamiento causal entre los distintos componentes del PCI (que en la

evaluación se denominó mecanismos de transmisión) fue analizado a partir de metodología cualitativa con entrevistas, grupos focales y documentos del diseño del PCI. Esto significa que se satisficieron los estándares de evaluabilidad y de lógica de la intervención tanto en la fase de diseño como de ejecución de la evaluación correspondientemente.

El estándar de coordinación y alineamiento, al igual que los dos anteriores, entra a jugar un papel importante en la etapa de diseño de la evaluación. Se toma en cuenta que el PCI se desprende del Plan Nacional de Desarrollo del 2010 “Prosperidad para Todos”, como una política de agua, saneamiento e higiene, la cual tiene dentro de sus fines el aumento en la calidad del agua. Sin embargo, la calidad del agua no es un objetivo del PCI, puesto que este se dirige a solucionar problemas de conectividad y acceso.

Esta fue la principal discusión que se dio entre el DNP, el evaluador/consultor y MVCT (lo cual se presentará a profundidad en la sección dedicada al análisis de la coordinación), es relevante decirlo en este punto, porque el hecho de considerar medir el PCI a partir de sus impactos en la calidad del agua en efecto, desconoció el alcance del PCI y su relación con otros programas (nacionales, departamentales y locales), los cuales todos en conjunto se dirigen a modificar la calidad del agua. Esto significa que el estándar de coordinación y alineamiento fue parcialmente satisfecho.

En el Pliego de Condiciones Definitivo hay un total de 24 preguntas que direccionan la evaluación de impacto y 8 preguntas para la evaluación de diseño del PCI (DSEPP-DNP, 2014c, pp. 54-55). Hay una adecuada correspondencia entre los objetivos específicos (p. 41) y las preguntas de evaluación con las cuales se busca alcanzarlos. Adicionalmente, de los criterios de evaluación aplicados por la OCDE, las preguntas direccionan a las siguientes preguntas:

- Eficiencia: “¿Los cambios que experimentan los beneficiarios del Programa en las variables de interés son relevantes y suficientes para justificar la intervención?”.
- Efectividad: “¿Cuáles son las principales estadísticas descriptivas de las variables de resultado y de control utilizadas en la evaluación?”.
- Impacto: “¿Cuáles son los impactos del Programa en las familias beneficiadas en relación con el mejoramiento de su salud por la disminución de episodios de diarrea y/o enfermedades asociadas al mal almacenamiento y uso de agua al interior de la vivienda?”.

Esto significa que los estándares correspondientes a cuestiones a responder y selección y aplicación de criterios de evaluación fueron completamente satisfechos al interior del proceso de evaluación. Por último, en la etapa de diseño de la evaluación, la selección de la metodología (recolección y análisis de datos) estuvo en correspondencia con el enfoque mixto entre evaluación de impacto y evaluación de diseño. Esta metodología es seleccionada por el Comité de Diseño, queda enmarcada en el Pliego de Condiciones Definitivo y es desarrollada por parte del

evaluador consultor en su primer producto, y diferencia entre distintos niveles de resultados. Esto hace que el estándar de selección de enfoque y método también haya sido satisfecho.

Con respecto a aquellos estándares que funcionan conceptualmente como componentes del estándar de informe de la evaluación[54], para su medición se procedió a su verificación utilizando los productos de la consultoría. Dentro de estos productos no hay una presentación puntual del contexto de la intervención, sin embargo, a lo largo del informe final se describen aspectos de este como el contexto de políticas, el contexto de desarrollo, el contexto institucional y las relaciones y participación de los actores implicados, por lo que el estándar de contexto de la intervención fue satisfecho.

Con respecto a la claridad del análisis, en efecto, en el informe se presentan de manera separada los hallazgos y las conclusiones, y en la ficha técnica y el PTIR se presentan por separado y de manera clara las conclusiones y las recomendaciones de la evaluación. Por esta razón se evidencia una línea clara de razonamiento entre los hallazgos, las conclusiones y las recomendaciones. De particular importancia para este punto, es la inclusión de una “matriz de respuesta a preguntas orientadoras” dentro del informe final (CNC, 2015b, pp. 156-158) en donde se señala la parte del documento en donde se responde a cada una de las preguntas de evaluación. Esto satisface todos los componentes de los estándares de claridad del análisis y de respuesta a preguntas evaluativas.

En el informe final también se realiza una adecuada explicación de la metodología de evaluación, esta se divide en metodología para la evaluación de impacto, metodología para la evaluación de diseño del PCI y el diseño muestral cuantitativo y cualitativo (pp. 8-22). También se realiza un análisis costo-beneficio, cuya metodología también es explicada suficientemente (p. 87). Son muy escasas las menciones que se hacen de los límites y deficiencias de la metodología, sin embargo, existen (y se presentan un poco más adelante), por lo que el estándar de explicación de la metodología en el informe final también fue satisfecho completamente.

La muestra cuantitativa seleccionó con muestreo aleatorio simple a hogares intervenidos y no intervenidos de los barrios en los que se implementó el PCI en 18 municipios (Apartadó, Armenia, Barrancabermeja, Barranquilla, Buenaventura, Cartagena, Florencia, Girardot, Ibagué, Leticia, Montería, Neiva, Pasto, Quibdó,

[54] Esto son: explicación de la metodología, lógica de la intervención, validez y fiabilidad de fuentes de información, mención de discrepancias entre miembros del equipo evaluador, Mención de los cambios y limitaciones de la evaluación, contexto de la intervención, y respuesta a las preguntas evaluativas.

Riohacha, Santa Marta, Sincelejo y Valledupar). En total fueron 1816 hogares encuestados para el grupo de tratamiento; 1553 hogares para el grupo de control (p. 17) y 208 encuestas a juntas de acción comunal. La muestra cualitativa consistió en 28 entrevistas a funcionarios de MVCT, empresas de servicios públicos, alcaldías e interventorías de las obras, y 36 grupos focales con 276 beneficiarios y no beneficiarios (p. 19). El análisis cualitativo ayudó a validar el análisis cuantitativo. Con todo esto, los estándares de validez y fiabilidad de fuentes de información y consulta y protección de implicados fueron satisfechos.

Dentro del documento del informe final no hay un resumen ejecutivo. Este resumen, por lo general, es presentado como otro producto a parte que se entrega en conjunto con el informe final. En este caso, se realizó un resumen de 25 páginas que incluye la metodología, los hallazgos, las conclusiones y las recomendaciones de la evaluación. Lo anterior satisfizo el estándar de claridad y representatividad del resumen.

En el contrato de consultoría y en el pliego de condiciones definitivo se observan reglas claras en cuanto al cronograma de entrega de los productos y, frente al pago que se realiza por producto recibido a satisfacción dentro de las fechas previstas (DNP, 2014b, p. 6). A pesar de que los productos no tienen fecha de entrega, no se encuentra ninguna evidencia acerca de la activación de la cláusula penal compensatoria por incumplimiento parcial o total del consultor, por lo que la evaluación fue realizada dentro del plazo y el presupuesto asignado.

Ni en este informe final, ni en ninguno de los productos de la consultoría se encontró mención de discrepancias entre miembros del equipo evaluador. Teniendo en consideración el trabajo conjunto que se establece en el Comité de Seguimiento, en el cual se involucran a los usuarios para la interpretación de los hallazgos y la construcción de recomendaciones, es poco probable que se den este tipo de menciones. Esto significa que no se satisface el estándar relacionado con este punto.

Si existen menciones de los límites del proceso evaluativo, particularmente relacionados con la transformación a pesos o monetización de los logros obtenidos para el análisis de costo-beneficio (CNC, 2015b, p. 88), particularmente los relacionados con salud psicosocial y derechos sexuales y reproductivos (p. 107); también con la metodología cuasi-experimental de la evaluación de impacto, la cual denominan método de diferencias en diferencias[55], la cual, al no contar con

[55] La evaluación así describe su método cuasi-experimental: “Para poder aislar el real efecto de la intervención, la comparación entre GT y GC debe ser, en lo posible, de dos etapas: una primera en la que se compara a cada grupo consigo mismo en dos momentos en el tiempo (antes y después de la intervención), para ver cómo evolucionó cada grupo de manera independiente; y una segunda etapa en la que se compara la diferencia entre esas diferencias, es decir, se compara el cambio de un grupo con el cambio del otro grupo. A esto se le llama una doble diferencia –o, más técnicamente, el método de diferencias en diferencias” (CNC, 2015b, p. 9).

información de línea de base para ambos grupos, requirió implementar un técnica denominada Propensity Score Matching cuyos resultados se definen como “menos robustos” (p. 9). Esto satisfizo el estándar de mención de cambios y limitaciones de la evaluación.

3.2.4 Reglas de agregación y control

En el pliego de condiciones (DSEPP-DNP, 2014c, p. 5) y en el contrato de consultoría (DNP, 2014b, p. 4) se establece como obligación del DNP (DSEPP) ejercer el control en el cumplimiento del contrato y expedir el certificado de cumplimiento a satisfacción. En correspondencia con los lineamientos generales a todos los contratos desarrollados por la DSEPP, este control se realiza a partir de la revisión que los funcionarios de la DSEPP y los participantes del Comité de Seguimiento realizan a los productos del consultor-evaluador, quien deberá incluir los comentarios que se le hagan a sus productos (p. 52).

La DSEPP también aumenta su grado de control frente a la obtención de los resultados potenciales de la situación de acción a partir del establecimiento de mecanismos de cobertura del riesgo (clausula décima segunda) al interior del contrato de consultoría. Estos mecanismos consisten en el deber del consultor de garantizar el cumplimiento de sus obligaciones a favor del DNP a partir de los mecanismos señalados en el artículo 11 del Decreto 1510 de 2013 (contrato de seguro contenido en una póliza, patrimonio autónomo y garantía bancaria). De esta manera, en dado caso que no haya cumplimiento o calidad del servicio, el contratista deberá pagar el 30% del valor del contrato, y si no paga los salarios y las prestaciones sociales o indemnizaciones laborales, un 10% de este.

Adicionalmente la DSEPP aumenta su control en la obtención de resultados con la cláusula penal compensatoria (décima quinta), en la cual se establece que en dado caso que se declara incumplimiento total o parcial habrá una pena pecuniaria equivalente al 30% del valor del contrato; así mismo se establecen multas (clausula décima octava) en caso de mora o incumplimiento parcial en cualquiera de las obligaciones contraídas por el consultor:

> “… las partes acuerdan multas sucesivas equivalentes al dos por mil del valor del producto, informe o insumo incumplido, pendiente por entregar y/o cumplir, por cada día de atraso en el cumplimiento de las obligaciones pactadas. Las multas tendrán un tope máximo del 20% del valor total del contrato y podrán ser compensadas con los valores que le adeude el DNP al consultor” (DNP, 2014b, p. 9).

Se debe aclarar que el control que ejerce el Comité de Seguimiento a los productos a partir de las observaciones que realiza, no es un control directo. Este control está

mediado por el papel que juega la “secretaría técnica” ejercida por la DSEPP, y tal como se observó, en el contrato y en el pliego de condiciones es el DNP quien ejerce la competencia de control sobre los productos, y no el Comité de Seguimiento. En este sentido, los usuarios previstos solo pueden ejercer un control indirecto, el cual está supeditado a la “consolidación, depuración y tratamiento” que la DSEPP realiza a los comentarios de los miembros del Comité de Seguimiento (DSEPP-DNP, 2014c, p. 51).

Todas estas reglas de agregación y la no existencia de multas en el desenvolvimiento del contrato, permiten establecer que todos los componentes de los estándares de control de calidad y de inclusión de observaciones de los implicados fueron satisfechos en este proceso evaluativo. Tal como se observó en el capítulo anterior, este punto está en tensión con la independencia del evaluador.

3.2.5 Reglas de posición

En cuanto a la independencia del evaluador y al desarrollo de un proceso libre y abierto, se tiene que las reglas de control hicieron que siempre existiera el riesgo (para el evaluador) de que el DNP decida activar las cláusulas de mecanismos de cobertura de riesgo o de multas, en dado caso que considere que hay un incumplimiento o mora en la entrega de los productos. En segundo lugar, este riesgo se ve acrecentado por el plazo máximo de una semana que tiene el consultor para presentar los productos con los comentarios del Comité de Seguimiento incluidos (DNP, 2014b, p. 4).

El impacto negativo de estas reglas de control sobre la independencia del evaluador, se debería ver disminuido con las discusiones que se realizan entre el Comité de Seguimiento y el evaluador acerca de estos comentarios (y que surgen cuando existen marcadas diferencias en las posturas de los participantes). Sin embargo, el proceso de elaboración de los productos, que es descrito en los conceptos técnicos de estos, muestra que los comentarios fueron generalmente aceptados e incluidos por parte del equipo evaluador.

Más allá de incluir o no los comentarios (lo cual es deseable desde muchas perspectivas teóricas en evaluación), es necesario profundizar en la naturaleza de estos, para poder identificar si estuvieron direccionados a cambiar las valoraciones y recomendaciones del evaluador sin una adecuada justificación. Los siguientes son los comentarios más significativos realizados al informe final y al PTIR:

> “Se solicitó la aclaración de los resultados de las tablas 18, 19, 20, 21 y 30, debido a que el ejercicio de réplica realizado por el DNP no arrojó el mismo resultado por ellos presentado” (DSEPP-DNP, 2015a, p. 3).
>
> “En relación con las referencias que hacen a cobertura en acueducto y alcantarillado y calidad del agua, se solicitó señalar cuáles son los datos para

> los municipios visitados, ya que el Ministerio menciona que muchas de estas alusiones se hacen sin el soporte en cifras" (p.3).
>
> "Se solicitó aclarar en el documento a qué hacen referencia cuando señalan que muchos, algunos o pocos beneficiarios reportan algo, ya que algunos apartes del documento parecen contradictorios en los resultados que reportan" (p. 3).
>
> "Se requirió que se presenten explicaciones de todos los resultados que dan significativos" (p. 3).
>
> "Se reiteró la solicitud de incluir el N de los cálculos que se realizaron en las variables de interés" (p. 4).
>
> "Se solicitó puntualizar cada recomendación. A manera de ejemplo se remarcó la idea central de algunas de las recomendaciones que se mencionan para denotar que es la idea principal. Posterior a señalar concretamente la recomendación, se debe incluir el aparte de justificación, para lo cual solicitamos que se señalen puntualmente en qué página del informe encuentran el resultado que la fundamenta. Luego, deben mencionar cuál sería el mecanismo de implementación y finalmente los indicadores de seguimiento" (2015b, p. 3).
>
> "En relación con el principal resultado de la evaluación de disminución de la probabilidad de contraer EDA en niños menores de 5 año y un efecto menor en jóvenes, se solicitó considerar que la evaluación podría recomendar criterios de focalización del programa diferenciada para hogares con niños menores de 5 años o en Hogares Comunitarios de Bienestar o en instituciones educativas" (p. 3).

En general, estos comentarios no incidieron en las valoraciones y recomendaciones del evaluador, en el PTIR no se encuentra ninguna recomendación orientada a la focalización de la cual habla la última cita, por lo que a pesar de que se incluyeron la gran mayoría de los comentarios, el único que está direccionado a restringir la independencia del evaluador, no tuvo realmente impacto negativo. A pesar de haber un número considerable de disuasivos institucionales para que el consultor prefiriera no debatir o defender sus resultados en caso de inconformismo por parte del DNP o de MVCT, en el proceso de evaluación se satisficieron los estándares de independencia del evaluador, y del proceso evaluativo libre y abierto.

3.2.6 Reglas de información

En la etapa de diseño de la evaluación, el estándar que se puede ubicar dentro de las reglas de información es el de documentación relativa a la finalidad y las

expectativas. Los estudios previos, y posteriormente, el proyecto de pliego de condiciones contiene toda la información acerca de los acuerdos a los que llegaron la DSEPP y los miembros del Comité de Diseño de la evaluación. Este proyecto de pliego de condiciones se publicó para la apertura del proceso de contratación, fue estudiado a profundidad por parte de los proponentes, y estos realizaron comentarios y preguntas en la audiencia de aclaraciones, luego de la cual se establece el pliego de condiciones definitivo.

En su conjunto, para esta evaluación no existieron grandes diferencias entre el documento de proyecto de pliego de condiciones y el pliego de condiciones definitivo, a parte de algunas correcciones de redacción que se le realizaron al primero. En estos documentos, en efecto, se cuenta con todos los componentes del estándar de documentación relativa a la finalidad y a las expectativas.

En la etapa de ejecución de la evaluación existieron reglas claras acerca de la radicación de los informes y productos. Además de la responsabilidad expresa dentro del contrato y el pliego de condiciones, se especificaron los siguientes informes/productos (DSEPP-DNP, 2014c, pp. 61-65):

1. Informe de propuesta metodológica.
2. Informe de avance del trabajo de campo.
3. Informe final del trabajo de campo.
4. Informe de resultados de la evaluación (informe final).
5. Plan de Transferencia e Implementación de Recomendaciones (PTIR).

Todos estos informes fueron comunicados al Comité de Seguimiento, fueron comentados por sus miembros y fueron corregidos con base en dichos comentarios. Lo anterior demuestra que los productos del evaluador/consultor, incluyendo el informe final, fueron comprendidos por los usuarios de la evaluación, por lo que el estándar de informe de la evaluación (sin tener en consideración los otros estándares que funcionan como sus componentes) fue satisfecho también.

Al final del proceso, el informe final de la evaluación se difundió a través de por lo menos dos canales. En primer lugar, se difundió al interior de los miembros del Comité de Seguimiento y del equipo líder de la formulación e implementación del PCI del MVCT. En segundo lugar, para un público más amplio, el informe final fue y aun es difundido en el portal de internet de SINERGIA en conjunto con la ficha técnica de la evaluación y de su resumen ejecutivo. Lo anterior satisfizo todos los componentes del estándar de difusión de la evaluación.

3.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación del Programa de Conexiones Intradomiciliarias

Para este proceso de evaluación, las únicas entidades que participaron en el Comité de Diseño y en el Comité de Seguimiento fueron el DNP, FONADE y MVCT. Estas últimas dos entidades vinieron a participar como usuarios previstos, siendo esta última la más importante en la medida que es quien lidera la formulación e implementación de la política pública. Los entrevistados mostraron que tuvieron interés, apertura al aprendizaje y compromiso con la evaluación, así como la capacidad de tomar decisiones y de influir lo suficiente como para poder implementar las recomendaciones que surgieran de la evaluación. El equipo de trabajo conjunto se reglamentó en el pliego de condiciones definitivo en concordancia con las disposiciones relativas al Comité de Seguimiento del Decreto 1290 de 2014, tal y como se presentó en el capítulo anterior.

Sin embargo, a pesar del hallazgo e involucramiento de usuarios clave con suficiente capacidad, particularmente de MVCT, muchas de las recomendaciones del PTIR se construyeron para que fuesen implementadas por entidades que no participaron en el proceso de evaluación, pero con las que constantemente debe coordinarse MVCT para la implementación del PCI. Entre estas entidades están: Ministerio de Salud y Protección Social, empresas de servicios públicos (de agua), SENA y alcaldías. La no participación de estos actores afecta negativamente la apropiación de los resultados de la evaluación por parte de quienes deben implementarlos.

La dinámica de este equipo de trabajo conjunto presentó dos fallas principales. En primer lugar, al momento de revisar las opciones de preguntas de evaluación y en el trabajo conjunto sobre los usos de la evaluación, los usuarios de la evaluación manifestaron que hubo diferencias en los intereses que los participantes tenían de la evaluación, lo que resultó plasmándose al momento de acoplar el alcance de la evaluación al alcance del PCI:

> "… creemos, y lo voy a decir realmente, que la percepción que nosotros tuvimos es que ellos querían con esa evaluación medir otro tipo de cosas, es decir, aprovechar y medir otros factores que no eran específicamente del programa (...) ellos querían el tema de mejoramiento de vivienda, el tema de cuántas familias con menor de edad hay, bueno, todo ese tipo de cosas, que en las discusiones parte de eso era que se estaban saliendo del alcance que queríamos que se mirara del programa (...) y las discusiones se centraban era que nosotros les explicamos a la luz de la norma, cómo se estaba implementando y cuáles eran las limitaciones y restricciones, que en ese momento la norma lo estaba exponiendo de esa manera (...) y no podíamos

> en ese momento cambiar la norma, o interpretar por fuera de la norma las cosas para que se ajuste a la evaluación que se estaba haciendo en ese momento, esa fue parte de las discusiones" (Coordinador PCI, MVCT).
>
> "… cuál fue la dificultad que tuvimos con el equipo de evaluación: explicar el alcance del programa, el programa como es del Viceministerio de Agua y Saneamiento, tiene un objetivo principal y es el tema de conexión, lo que queremos garantizar es que una vivienda que tiene al frente la red de acueducto y alcantarillado quede conectada y ese es nuestro objetivo" (Coordinador PCI, MVCT).

Esto refleja una escasa atención o escucha (reacción) a las preocupaciones y demandas de los usuarios previstos por parte del equipo evaluador/consultor y de la DSEPP. Incluso se evidencia una fuerte reticencia y prevención hacia la determinación conjunta del alcance de la evaluación, la cual se observa como una injerencia en la independencia del evaluador:

> "…(a los evaluadores) si se les explicó el alcance del programa y todo ese tipo de cosas, pero de ahí, en las reuniones que hubo, si había como cierto choque y es como de alguna manera tratando de explicarles, de darles a entender el alcance del programa, la complejidad, cómo evaluarlo, tal vez desde que punto de vista verlo, pues había como ese choque, de alguna manera, que los evaluadores piensen que de alguna manera nosotros estuviéramos coaccionándolos para dar a unos resultados esperados que quisiéramos nosotros, pero no, esa fue realmente la percepción que hubo, y entonces se volvió de alguna manera como un choque" (coordinador PCI, MVCT).

Los desacuerdos alrededor del alcance del programa se extendieron durante todo el proceso de evaluación y alcanzaron las últimas etapas de la coordinación de interpretación de los hallazgos, particularmente alrededor del tema de calidad del agua y de construcción conjunta de las recomendaciones, en donde el evaluador/consultor sugirió la ampliación de la instalación de enchapes y de tanques de agua sin considerar el alcance del programa:

> "La gran discusión que tuvimos fue el tema de la calidad del agua (…) estaban, de alguna manera, tirándole la responsabilidad al Programa de Intradomiciliarias, que debía garantizar o mejorar la calidad del agua, entonces un tema de calidad del agua es el tratamiento que tu le haces, y el programa no está encaminado a mejorar calidad sino a mejorar conectividad, tú al instalar un inodoro o al instalar un lavaplatos, no estas instalando un filtro para el tratamiento del agua ni nada de eso (...) por más que nosotros tratamos de contextualizar, ellos de alguna manera lo dejaron (…) y colocaron (…) en el tema de calidad del agua, no generó ningún mejoramiento o impacto, a pesar de que el objetivo del programa no era ese" (Coordinador PCI, MVCT).

> "… es que el concepto de ellos no era el enchape sobre las obras que había llevado el programa, sino si se podía entrar a una casa exclusivamente a hacer el enchape, y eso el programa no podía llegar porque no era mejoramiento, eso fue una controversia muy fuerte porque ellos lo miraban como parte de salubridad (…) la casa tiene todo, tenía las redes, pero el baño requería el enchape y eso no lo puede hacer el programa, porque no era mejoramiento. Eso se les dijo a ellos, se les informó que no se podía, pero lo dejaron ahí" (Contratista MVCT).

> "…el tema de tanques el que también es una recomendación general de que todas las intervenciones deben ir con tanque prácticamente es lo mismo, el tema de tanques es un tema costoso (…) por lo que ya se está comiendo gran parte del subsidio (…) el tanque si lo instalamos, pero lo tenemos limitado y condicionado a cuando tenemos problemas de continuidad en los barrios que vamos a intervenir" (Coordinador PCI, MVCT).

En segundo lugar, los entrevistados manifestaron inconformismo frente a su escasa participación en el trabajo de campo. Se presentaron situaciones en que, a pesar de manifestar su interés y tener experiencia en recolección de datos en las regiones, la DSEPP les informaba con muy poco tiempo de anticipación de las fechas de recolección de datos, por lo que no podían asistir (contratista MVCT). Este punto es relevante porque afecto directamente la calidad de los datos recolectados:

> "… lo que detectamos es que las personas que aplicaban las encuestas, eran experimentadas pero no en este tipo de encuestas tan técnicas, y yo me acuerdo de una casa de una señora en Girardot que no entendió ni una sola pregunta que le hicieron, entonces ellas si son expertas en la técnica de la aplicación de encuestas, pero yo creo que para estos casos técnicos si habría que tener más interpretación y unirse más al Ministerio o al ejecutor (...) se requería un mejor entrenamiento por parte del equipo evaluador (...) más orientado a conocer el programa" (contratista MVCT).

> "… resulta que este programa necesita ser muy entendido por cada una de las personas que van a participar y que van a estar empapadas con la comunidad, y por eso ese si fue un tema importante, como no vino el personal acá que iba a evaluar en campo, sino los directivos, ahí pensamos nosotros, que también el canal de comunicación se nos fue distorsionando y se nos fue distanciando" (Coordinador PCI, MVCT).

La distorsión en la comunicación es explicada por parte de los entrevistados a razón de que el trabajo de campo quedó a la luz de la interpretación de la normatividad y del programa que tenían quienes recolectaron los datos (coordinador PCI, MVCT). Estas discusiones sobre el alcance y la recolección de datos reflejan que, dentro de los debates metodológicos, ni la DSEPP ni el evaluador/consultor estuvieron

siempre en sintonía con las preocupaciones de los usuarios y permitieron que sus intereses dominaran el proceso de ajuste de prioridades.

Las anteriores citas también son dicientes frente a los debates metodológicos y la adaptación de la metodología como componentes de la coordinación, puesto que evidencia fuertes falencias en la capacidad de adaptar el diseño de la evaluación en respuesta a las necesidades cambiantes de los usuarios. Así mismo, tal como se sugirió en el capítulo anterior, al comparar el proyecto de pliego de condiciones y el pliego de condiciones definitivo (cuya transición está mediada por la audiencia de aclaraciones en donde los postulantes pueden proponer recomendaciones a este proyecto de pliego de condiciones, y, por lo tanto, a la metodología seleccionada entre la DSEPP y los usuarios previstos), se evidencia que no existe ninguna diferencia, por lo que ninguna sugerencia de modificación de la metodología de ningún postulante (incluido el evaluador/consultor seleccionado) fue tenida en cuenta por el DNP.

> "Este tema del esquema anticipado de pedidos al contratista (…) no lo hemos utilizado, porque por la forma de implementación del programa es muy difícil aplicarlo, (…) en el mes 1 el municipio hace el levantamiento de la encuesta, pero de ahí tenemos varios meses para estructurar el proyecto, que lo hacemos a través de nuestro ejecutor que es FONADE, y ahí va pasando el tiempo, eso que significa que las cosas en el barrio van cambiando, entonces una casa en la que se identificó que necesitaban los cinco aparatos, pues cuando llegamos con el contratista ya necesitan solamente dos, (…) el contratista no puede entrar desde un principio a pedir a los proveedores: necesito tantos inodoros, necesito tantos lavamanos, porque la situación en campo me va a cambiar" (Coordinador PCI, MVCT).

Tal como se observa, por la misma institucionalidad del PCI, su división de tareas y sus tiempos, esta recomendación no es factible de ser implementada. Al preguntar ¿por qué terminaron recomendando eso si realmente no era factible o posible? Se respondió:

> "… tenemos que reconocer que tuvimos varias reuniones en las cuales hubo varios encuentros de opiniones y desacuerdos, en que ese fue uno de tantos, porque realmente las personas que fueron a región a tomar esas muestras no conocían el programa, entonces cuando yo no conozco, pues me confío en que todo lo que me dicen es verdad, y ahí hay una primera falencia, nosotros les dijimos al DNP y al CNC que no compartimos esa recomendación, pero nunca la cambiaron" (Coordinador PCI).

Esto significa que, a pesar de que se incluyeron los comentarios del Comité de Seguimiento presentados en la sección de reglas de posición, en temas centrales, existieron fuertes fallas en el involucramiento de los usuarios en la construcción de recomendaciones y en la interpretación de los hallazgos. Por último, quedan los temas de la planeación conjunta del seguimiento a la implementación de los hallazgos de la evaluación y la búsqueda proactiva de uso. Esta última, como se vio

en el capítulo anterior, en dado caso que se realice, sería por acción de la DSEPP puesto que el alcance de los contratos rompe la relación evaluador-usuarios previstos al momento de hacer entrega de los últimos productos corregidos.

La única información que se encontró alrededor de un seguimiento a la implementación de los hallazgos de la evaluación está plasmada en los indicadores que se construyen en el PTIR para observar el grado en el que se implementan las recomendaciones, estos indicadores son:

- Disminución en casos de EDA (Enfermedades Diarreicas Agudas).
- Evaluación del programa utilizado en RISP (Registro Individual de Prestación de Servicios).
- Índice de continuidad del servicio de agua.
- Presión del servicio de agua.
- Baños acabados “enchapados”.
- Número de programas de capacitación realizados por el SENA.
- Porcentaje de personas contratadas por el PCI que fueron capacitadas por el SENA.
- Suficiencia de proveedores.
- Porcentaje de procesos de selección en los que se exigió al contratista entrar en consorcio con un proveedor de materiales.
- Aplicación de encuesta a población beneficiaria.
- Tabulación y evaluación de resultados.
- Mujeres capacitadas en Derechos Sexuales y Reproductivos.
- Participación de las mujeres en el PCI.
- Reducción de morosidad.
- Socializaciones acompañadas por el MVCT.
- Folletos entregados a la población.
- Capacitaciones realizadas.
- Acompañamiento de actores locales a la intervención.

A pesar de que la DSEPP respondió afirmativamente acerca de la realización de seguimiento (no sistemático) a la implementación de las recomendaciones y de los hallazgos de la evaluación en general, la información recolectada con las entrevistas evidencia que este seguimiento no es conjunto con los usuarios de las evaluaciones, y, por lo tanto, no constituye un escenario de coordinación. Al preguntar acerca de si el DNP les ha preguntado alguna vez cómo va la implementación de estos resultados, los funcionarios de MVCT contestaron:

> “(entrevistado): No, la verdad es que, no. Si las tenemos muy presentes porque pues el programa realmente ha tenido buenos resultados y hemos tratado de irlo optimizando y mejorarlo poco a poco.
> (entrevistador): Pero ¿se terminó cuando les entregaron los productos?
> (entrevistado): Sí, hasta ahí. Lo último fue cuando nos entregaron el documento y ahí no volvimos a tener contacto.” (Coordinador PCI, MVCT).

Esto quiere decir que los anteriores indicadores nunca han sido aplicados, ni por el evaluador, ni por el DNP. Tampoco han funcionado como una herramienta que fomente o fortalezca el uso de la evaluación. En este sentido, hay una total carencia de coordinación para el seguimiento a la implementación de recomendaciones y de una búsqueda proactiva del uso por parte de los encargados de la evaluación.

3.4 Análisis de la efectividad de la evaluación del Programa de Conexiones Intradomiciliarias.

En esta sección se presentará cada una de las recomendaciones seguido por una descripción de su estado de implementación por parte de MVCT. Para esta medición se utilizaron datos provenientes del PTIR y de entrevistas a los funcionarios de esta entidad. La primera recomendación se trata de que MVCT haga un seguimiento a indicadores de salud (CNC, 2015c, p. 5) que son afectados por el PCI, de tal manera que se puedan verificar sus impactos. Para esto se sugiere la utilización del Registro Individual de Prestación de Servicios (RIPS) que está a cargo del Ministerio de Salud y Protección Social. En cuanto a su implementación se comenta:

> "… digamos, que de ahí acá, no hemos todavía implementado ese tema, o sea estamos como, incluso el año pasado, empezamos conversaciones con el Ministerio de Salud para ver cómo podemos ir haciendo ese seguimiento en los proyectos que vayamos implementando, pero, pues la realidad ahorita, es que posterior a eso, no hicimos como otra medición, respecto al tema de salud, no la hemos hecho (…) un avance es que ya tuvimos conversaciones con el Ministerio de Salud, y ellos tienen información de línea base, digamos de, las Regiones o los Municipios, entonces la idea es, con ellos, esperamos escoger ahorita este año que ya tengamos, digamos, recursos para el programa, coger un proyecto de esos piloto, en conjunto con el Ministerio de salud, y hacerle una medición" (Coordinador PCI, MVCT).

Como se observa, la recomendación no ha sido implementada, y a pesar de que se contó con el PTIR corregido en septiembre de 2015, hasta ahora han iniciado conversaciones con el Ministerio de Salud para hacer un seguimiento en consonancia con esta recomendación. Se debe resaltar que este seguimiento está pensado hacerse para nuevos proyectos que se realicen dentro del PCI, pero no para verificar los impactos que se midieron con la evaluación, lo cual era la intención de esta recomendación.

La segunda recomendación se dirige a verificar el requisito de calidad del servicio (CNC, 2015c, p. 8), en donde MVCT debería exigirles a las empresas de servicios públicos que realicen la verificación del cumplimiento de continuidad y presión del servicio de agua, esto con la intención de identificar los barrios que necesitan instalación de tanques de agua. Adicional a los problemas de factibilidad por el costo

que esto puede implicar, lo cual se presentó en la sección anterior, esta recomendación tiene la grave falencia de ser una actividad que ya había implementado MVCT, incluso antes de que terminara la evaluación:

> "En ese tema, digamos que, nosotros sí, sí hemos implementado como unas actividades relacionadas con la confirmación de esos aspectos (…) Pues la realidad es que es algo que identificamos ya antes, es algo que identificamos ya. Lo que pasa es que como ellos nos certifican, digamos, documentalmente: en papel, pero cuando ya vamos a campo a realizar el trabajo, es ahí donde uno va adquiriendo, digamos, mucha experiencia de las condiciones reales de prestación de los servicios en ese tipo de barrios." (Coordinador PCI, MVCT).

La implementación de esta recomendación al no ser el resultado de los hallazgos de la evaluación, y gracias a que funciona a modo de "documentación" o como certificación de lo que ya se estaba implementando en MVCT, es un ejemplo de justificación, argumentación o legitimación de decisiones ya tomadas. En efecto, si es algo que ya había considerado el MVCT, entonces el modelo mental de los usuarios previstos no sufre ningún cambio a causa de la recomendación, por lo que no se mejora la toma de decisión y, por ende, no sirve para decir que la evaluación fue utilizada o que el proceso de evaluación fue efectivo en este punto.

La tercera recomendación se dirige a la adecuación estética de los baños, en donde MVCT debería incluir en el manual operativo para los contratistas indicaciones acerca de la prioridad al acabado (enchape) de la unidad sanitaria básica de los hogares (CNC, 2015c, p. 12). Como se observó en la sección anterior, esta recomendación reflejó problemas alrededor de determinar el objetivo y el alcance del programa y, por ende, de la justificación del gasto:

> "… ese fue otro tema que tuvimos la discusión con ellos, que a la vez se refleja en la dificultad que hay para poder explicar el alcance del programa (…) tenemos entonces unos topes de subsidio que son muy justos para lo que hay que hacer, y también la normativa del programa está muy encaminada a que debo instalar esos 5 aparatos y esas redes, por lo tanto los otros temas, como por ejemplo el de enchapes, muchas veces nosotros lo tenemos controlado porque no podemos expandir ese tema de enchapes porque en primer lugar son más costos, y tenemos distintas situaciones en donde por ejemplo si gastamos los recursos del subsidio en enchapar un baño, no nos queda plata para instalarle el lavadero y no nos queda plata para instalarle el lavaplatos que es el objetivo del programa, entonces por eso nosotros ese tema de los enchapes lo tenemos muy limitado exclusivamente a la zona húmeda de la ducha, que es un área mínima ahí con unos costos mínimos que no nos afectan mucho la esencia del programa" (Coordinador PCI, MVCT).

Con esta recomendación, puede evidenciarse que a pesar del trabajo conjunto entre el equipo evaluador y el Comité de Seguimiento, pueden surgir recomendaciones que son descontextualizadas y por ende no factibles de ser implementadas, en este caso, por la limitación de recursos establecidos por la ley y que tiene una particular priorización. La recomendación, en consecuencia, no fue implementada.

La cuarta recomendación sugiere capacitar a la población de los barrios beneficiarios en oficios que son demandados por el PCI (CNC, 2015c, p. 14). La forma de implementación debería pasar por identificar las profesiones y oficios que son requeridos por el programa, determinar las ciudades a intervenir que podrían presentar problemas de escases de personas entrenadas en los oficios requeridos, y luego establecer un convenio con el SENA para que realice las capacitaciones necesarias en los barrios. Respecto a su implementación, se manifestó:

> "… antes de que nos hicieran esa evaluación, nosotros ya teníamos un programa que se llama *Plan de Gestión Social* en donde se involucra a la comunidad en cada una de las etapas del programa (…) para que, las personas que sean beneficiadas con el programa, aprendan a utilizar los aparatos, cómo se hace el mantenimiento, todo ese tipo de cosas (…) para vincularlos al tema de que sean parte, como tal, de la ejecución de las obras, eso sí lo tenemos incluido, FONADE que es nuestro ejecutor, él tiene incluso una cláusula con el contratista de obra, en donde le obliga a que un porcentaje del personal que contrate el contratista, así sea, digamos, en mano de obra no calificada, el porcentaje que exige FONADE de personas de la comunidad es del 20 por ciento (…) también presente y estamos tratando, digamos, de articularnos con el SENA, o con otras entidades para que, ese personal, digamos, esté como preparado antes de que lleguen las obras" (Coordinador PCI, MVCT).

Al igual que con la anterior, a pesar de que se han implementado acciones en correspondencia con esta recomendación, se observa que estas fueron implementadas antes de la terminación de la evaluación y como producto de la misma dinámica interna de la política pública, por lo que no significa un cambio en el modelo mental de los usuarios, y no es un ejemplo de uso o de efectividad de la evaluación, sino de justificación de decisiones ya tomadas. La única parte novedosa de la recomendación, relativa al establecimiento de un convenio con el SENA, aun no se ha implementado.

En quinto lugar, se recomendó exigir al contratista que entre en consorcio con un proveedor de insumos. En este caso, cuando FONADE considere que en determinadas ciudades no hay proveedores con la capacidad suficiente, deberá exigir a los contratistas que se presenten al proceso de selección en consorcio con un proveedor de insumos, para garantizar el suministro de materiales necesarios (CNC, 2015c, p. 17). Al preguntar por esta recomendación se contestó algo muy similar a su opinión ante el esquema anticipado de pedidos:

> "… no era como muy conveniente de aplicarla, por la forma de funcionamiento de ese programa; por el tema de tiempos, porque no podemos casi que desde un tiempo anticipado determinar qué tantos insumos vamos a utilizar ¿si? sino que eso ya los identificamos más es en campo, cuando el contratista de obra ya está trabajando. Entonces ahí corremos el riesgo de que si, vinculamos, no sé, a un proveedor y previo a la contratación decimos que vamos a, no sé, a implementar o a utilizar ciertos aparatos, o cierta cantidad, digamos de, accesorios, tuberías, etc., y si eso no se utiliza en su totalidad, entonces ahí podemos tener, digamos, un problema con el contratista, porque nosotros al contratista le pagamos es por obra terminada y recibida satisfacción" (Coordinador PCI, MVCT).

Esta recomendación no fue implementada, y el hecho de que no sea factible de ser implementada significa que hubo problemas en el trabajo conjunto con los usuarios previstos. En sexto lugar, se recomendó realizar una encuesta de seguimiento a la población beneficiaria para poder evaluar variables que se reflejan luego de la implementación del PCI (como la satisfacción de los beneficiarios y la calidad de las obras). En el mismo PTIR aparece "La propuesta, que inclusive ya está siendo adoptada por el MVCT en algunas ciudades, es la realización de una encuesta de satisfacción de los beneficiarios del PCI un tiempo después de que han terminado las obras y las capacitaciones" (CNC, 2015c, p. 20). Lo cual es otro ejemplo de justificación de decisión ya tomada, y no de mejoramiento del proceso de toma de decisiones.

La séptima recomendación fue la inclusión de los Derechos Sexuales y Reproductivos en las capacitaciones que se le brindan a la población beneficiaria, para lo que se deberían incluir lineamientos al respecto dentro de los manuales de estas capacitaciones que MVCT les brinda a las alcaldías y a las empresas de servicios públicos (p. 24). De esta recomendación se comentó:

> "Bueno, ese es un tema, que sí, digamos, se identificó a través de la evaluación del DNP, ese es un valor agregado que no lo teníamos nosotros identificado, no lo teníamos, sin entrar a ver el programa, y gracias a la evaluación del DNP fue que se identificó ese tema. Que es algo que, digamos, lo hemos tenido presente también, pero la realidad es que estamos viendo cómo poderlo trabajar o desarrollar mejor" (Coordinador PCI, MVCT).

La recomendación propone algo novedoso, y a pesar de la pertinencia que encuentran los funcionarios del MVCT, y de que en el PTIR se pretende realizar una explicación de su modo de implementación, a la fecha es algo que sigue sin implementarse. Además, si bien es algo que parece que puede implementarse en un futuro próximo, hasta el momento eso es algo que aún no se ha realizado.

En octavo lugar se recomendó facilitar las formas de pago del servicio de agua a partir de una articulación con las empresas de servicios públicos, en donde se deben

contemplar distintas estrategias (como financiación de deudas a morosos) para adaptarse a la volatilidad de los ingresos de las familias beneficiarias del PCI (CNC, 2015c, p. 26).

> "Bueno, ese es un tema que si lo hemos identificado en campo y hemos tratado de darle manejo con las empresas (…) En este sentido, el tema de que sea un deudor moroso o no, no es un criterio para excluir a la vivienda a la luz de la normativa vigente (…) entonces en ese sentido, nosotros trabajamos con la empresa de servicios, dejándole claro que ellos no pueden, de alguna manera, excluir a una vivienda por el hecho de que esta vivienda esté deudora en el servicio. Y esa ha sido, digamos, una de las situaciones que hemos tenido que irle dando manejo, en donde las empresas, ahí, precisamente, entran como a manejar acuerdos de pago" (Coordinador PCI, MVCT).

A pesar de que se dijo que el problema es una situación que han identificado en campo, no se implementaron acciones equivalentes antes de la evaluación, razón por la que si es una recomendación que se ha implementado gracias a la evaluación.

La novena recomendaciones es realizar un seguimiento y control a las socializaciones y capacitaciones realizadas por las empresas de servicios públicos y las alcaldías. Para esto, se sugiere que el MVCT haga parte del equipo de socialización del programa para asegurarse que la información entregada es adecuada (CNC, 2015c, p. 28). A pesar de poderse evidenciar que el MVCT ya venía implementando una socialización con su Plan de Gestión Social, se evidencia un fortalecimiento en el siguiente sentido:

> "...lo que hemos hecho es fortalecer un poco los canales de socialización en donde le damos un poco más de peso y responsabilidad al municipio, en donde ya en los convenios que firmamos con los municipios, nosotros exigimos que haya una coordinadora social 100% dedicada a ese tema, estamos fortaleciendo el tema de capacitaciones, ya, estamos haciendo capacitaciones a distintos grupos, entonces nos estamos concentrando en capacitar a todo el equipo de obra con todo su equipo social y técnico, a la interventoría, estamos capacitando también a municipios, a empresas de servicios, y también ya a raíz de eso, tenemos ya un programa social con una capacitaciones un poco más detalladas y más dedicadas a la población" (Coordinador PCI, MVCT).

Tal como se observa, el protagonismo de MVCT en las capacitaciones y la institucionalización de mayores exigencias a las alcaldías desde la firma de los convenios, ejemplifican la implementación de la recomendación. Esta implementación también ejemplifica un uso efectivo de la evaluación.

Por último, la décima recomendación se dirige a la interacción entre los actores y la sostenibilidad del programa. Para esto, MVCT deberá incorporar en el diseño del

PCI a las alcaldías y a las empresas de servicios públicos en cuanto al acompañamiento que estos actores deben realizar a la intervención (CNC, 2015c, p. 29). Ante esto se manifestó:

> "Bueno en ese tema, ese tema sí puedo decir que, que si bien, por el objetivo de la evaluación de DNP, lo incluyó ahí; si estaba contemplado desde un principio cuando se diseñó el programa. En la normativa del programa, para que un Municipio pueda ser beneficiado con el programa, a través de un financiamiento por parte del Ministerio de Vivienda, dice que, el Municipio para poder acceder a ese financiamiento debe aportar en una contrapartida, que esa contrapartida puede ser en: recursos económicos o puede ser en la implementación de toda la estrategia social y en el acompañamiento al proyecto conforme a los lineamientos que establezca el Ministerio; en ese sentido, el Ministerio tiene unos convenios de cooperación (…) ahí, el Ministerio establece y aclara cuáles son las obligaciones que debe cumplir el Municipio como contrapartida. Y ahí entre esas está que debe apoyar el Municipio en toda la estructuración del proyecto" (Coordinador PCI, MVCT).

La institucionalidad del programa ya incluía a los municipios y a las empresas de servicios públicos al interior del diseño de los proyectos del PCI, por lo que esta recomendación dice algo que ya se estaba implementado, no significó un cambio en el modelo mental del usuario, justifica una decisión previa, y no puede ser considerada como un caso de uso de la información producida por la evaluación.

3.5 Conclusiones

En este proceso de evaluación la mayoría de los estándares de calidad fueron satisfechos a partir de reglas y de la interacción de los actores en cumplimiento de estas. Sin embargo, se presentaron falencias en cuanto a las reglas de alcance y resultados potenciales, en donde no se satisfizo el estándar de oportunidad, pertinencia y utilización. Si bien existen reglas transversales a todos los procesos de evaluación dirigidas a satisfacer esto, la divergencia de intereses entre el DNP y MVCT derivó en que el evaluador/consultor no dirigiera su acción principalmente al abordaje de la demanda de los usuarios previstos, sino a las demandas del DNP.

Con respecto a las reglas de frontera y los participantes, los estándares de participación de actores relevantes, enfoque asociativo y consideración de evaluación conjunta no fueron satisfechos, esto es así por la no inclusión de actores relevantes con los que MVCT se coordina para implementar el PCI. En reglas de elección, la única acción que no se cumplió fue la mención de discrepancias entre miembros del equipo evaluador al interior del informe final de la evaluación.

El análisis de coordinación evidencia que se hallaron e involucraron a los usuarios previstos de la evaluación, particularmente de MVCT y de FONADE, sin embargo, se dejaron de incluir otros actores clave como: Ministerio de Salud y Protección Social, empresas de servicios públicos (de agua), SENA y alcaldías, esto incidió negativamente en la apropiación que deben hacer estos usuarios de la evaluación. Tal como se observó, cuatro años después de contar con los resultados de la evaluación, las recomendaciones que se dirigieron a estos actores no se han implementado.

DSEPP, FONADE, MVCT y el equipo evaluador/consultor formaron un equipo de trabajo conjunto en el Comité Técnico de Seguimiento, el cual se reunió con regularidad, buscó abordar las necesidades establecidas en el Comité de Diseño de la evaluación y sirvió como escenario principal de coordinación para la evaluación. A pesar de esto, los datos de las entrevistas evidencian que los actores tuvieron diferentes intereses en cuanto a las preguntas y variables a medir en la evaluación. Este problema derivó en una serie continua de discusiones alrededor del alcance del PCI en todo el proceso de evaluación, lo que significó un muy bajo grado de escucha y reacción a las preocupaciones y demandas de los usuarios previstos por parte del evaluador/consultor y de la DSEPP; y una fuerte reticencia y prevención hacia la determinación conjunta del alcance de la evaluación.

Se encontraron falencias en la capacidad de adaptar el diseño de la evaluación en respuesta a las necesidades cambiantes de los usuarios, como se reflejó en su escasa participación en el trabajo de campo a pesar de su interés de hacerlo. Tampoco se incluyó ninguna sugerencia de modificación de la metodología de ningún postulante (incluido el evaluador/consultor seleccionado), fue tenida en cuenta por el DNP. Esto muestra que en general, el diseño de la evaluación fue bastante rígido y no se dieron discusiones para su transformación.

Estos desacuerdos alrededor de aspectos fundamentales de la evaluación se extendieron hasta las últimas etapas del proceso. No hubo un involucramiento constante de los usuarios previstos en la interpretación de los hallazgos de la evaluación (como se observó frente al tema de calidad del agua) y en la construcción de las recomendaciones (lo que se evidencia en diversas recomendaciones que no era factibles de implementar). Esto no quiere decir que el equipo de trabajo conjunto nunca direccionara su acción a satisfacer las necesidades de los usuarios previstos, pero se puede decir que, por lo menos, en muchos momentos de la coordinación, ni la DSEPP ni el evaluador/consultor estuvieron en sintonía con las preocupaciones de los usuarios y permitieron que sus intereses dominaran el proceso de ajuste de prioridades.

En las últimas discusiones que se dieron entre el Comité de Seguimiento y el evaluador/consultor, todos estuvieron de acuerdo en la formulación de un conjunto de indicadores que permitiesen evidenciar el avance en la implementación de los hallazgos de la evaluación. Pero al ser este el único escenario en el que se reflejó una planeación conjunta frente al seguimiento a esta implementación, se perdió la oportunidad de generar un espacio que potencializara el uso de la evaluación.

Los problemas evidenciados anteriormente están en coherencia con el bajo grado de efectividad que ha tenido el proceso de evaluación del PCI. De diez recomendaciones, sólo dos fueron implementadas, cuatro no fueron implementadas y otras cuatro funcionaron como justificación de decisiones ya tomadas. Estas últimas, para los usuarios previstos tienen la utilidad de servir a modo de documentación o soporte técnico de lo que se viene haciendo, pero independientemente de esta postura, el objetivo del proceso de evaluaciones es claro: mejorar la toma de decisión, lo cual no existe si la decisión es previa a la evaluación.

Esta evaluación fue realizada en un contexto en el cual había un fuerte interés por reformular el programa. Como se pudo observar en el análisis de efectividad, algunos de los mecanismos institucionales por medio de los cuales se ha pretendido implementar las recomendaciones incluye convenios con entidades del orden nacional, con empresas de servicios públicos y con municipios, además de los contratos a través de los cuales se implementa el programa. Adicional a esto, los últimos resultados de la evaluación se tuvieron disponibles el 28 de octubre de 2015 (DSEPP-DNP, 2015b, p. 3), a lo cual le siguió todo el proceso de elaboración de la Guía Metodológica del Programa de Conexiones Intradomiciliarias (2018), que adoptada como normatividad técnica con la Resolución 0528 de julio de 2018 del MVCT. Todo esto refleja que hubo apertura hacia el cambio en la institucionalidad del programa por parte de los usuarios previstos de la evaluación.

4. Efectividad de la evaluación institucional y de resultados del programa Planes Departamentales de Agua del Ministerio de Vivienda, Ciudad y Territorio

4.1 Presentación del programa Planes Departamentales de Agua

El programa Planes Departamentales de Agua (PDA) es un conjunto de estrategias de planeación y de coordinación interinstitucional, que buscan la armonización integral de los recursos y la implementación de esquemas eficientes para la prestación de los servicios públicos domiciliarios de agua potable y saneamiento básico, esto teniendo en cuenta las características locales, la capacidad institucional de entes territoriales y de las organizaciones prestadoras de los servicios públicos (Economía Urbana e Infométrika, 2015a, p. 34).

Dentro de sus componentes se encuentran: la facilitación de la coordinación interinstitucional entre sectores y entre niveles de gobierno; la modernización empresarial del sector de agua en todo el territorio nacional; el aprovechamiento de economías de escala mediante la estructuración de esquemas regionales de prestación; la articulación de diferentes fuentes de recursos junto con la facilitación del acceso del sector a crédito; mejorar el control sobre recursos y cumplimiento de la regulación y contar con planes de inversión con perspectiva regional (CONPES 3463 de 2007, p. 5).

El marco normativo del programa PDA incluye: CONPES 091 de 2015, CONPES 3464 de 2007, Ley 1151 de 2007, Decreto 3200 de 2008, Decreto 2371 de 2009, Decreto 2311 de 2010, Ley 1450 de 2011, Decreto 4214 de 2011, Decreto 2246 de 2012, Resolución MVCT 379 de 2012, Resolución MVCT 504 de 2013, Decreto 475 de 2015, y Ley 1753 de 2015. Además de las directrices incluidas en los planes de desarrollo a nivel nacional de los gobiernos desde 2006 hasta 2018.

Los actores que se buscan coordinar en el programa son: el MVCT, los entes territoriales, las consultorías, Autoridades Ambientales Regionales, Secretarías de Salud Departamentales, Oficinas de Registro de Instrumentos Públicos, Agencia

Nacional de Infraestructura, Findeter, Banca multilateral, empresas de servicios públicos (ESP), gestores PDA (que pueden ser gobernaciones o ESP), Consorcio Fiduciario FIA, Contratistas, usuarios.

4.2 Caracterización de las situaciones de acción del proceso de evaluación del Programa Planes Departamentales de Agua

4.2.1 Reglas de alcance y resultados potenciales

La evaluación al PDA se realizó con la finalidad de "fortalecer las capacidades del Estado y mejorar el proceso de diseño e implementación de la política sectorial" (DSEPP-DNP, 2014d, p. 3). Su relevancia radica en la necesidad de saber -luego de cuatro años de implementación- si el programa ha logrado la armonización de recursos del sector y el mejoramiento gradual de la implementación de los esquemas eficientes y sostenibles para la prestación de servicios de agua y saneamiento básico. Del mismo modo, se buscó que la evaluación permitiera conocer si el fortalecimiento de los procesos de planeación, asistencia técnica y control tuvieron incidencia en los indicadores de desarrollo institucional y de prestación del sector de agua (p. 3).

Se señaló como evidente la necesidad de contratar la evaluación del PDA, ya que: "…para el desarrollo económico y social del país representa una oportunidad muy importante conocer los avances del sector y realizar ajustes a la política que aseguren con mayor precisión el logro de los objetivos" (p. 4). En el mismo orden de ideas, en el informe final se plantea que la evaluación debe ser un medio para corregir los problemas que surgen durante la implementación de la política, por lo que su información "… debe ser incorporada en la toma de decisiones permitiendo realizar los ajustes pertinentes" (Economía Urbana e Infométrika, 2015a, p. 21).

Esto significó una adecuada aplicación de todos los componentes del estándar de fundamentación y con relación a la finalidad de la evaluación, se cumplieron en esta evaluación. Adicionalmente, la evaluación al tener por finalidad el aumento de la capacidad institucional del Estado (incluyendo la formulación de los indicadores para la evaluación), cumplió también con el estándar de desarrollo de capacidad.

En cuanto al objeto y alcance de la evaluación, esta se dirige a ser una mezcla entre una evaluación institucional y de resultados del PAD, esto quiere decir que el alcance de la evaluación se direccionó a identificar la capacidad institucional de los actores para cumplir con los objetivos del PDA y también el grado de cumplimiento de los objetivos y metas del programa (DSEPP-DNP, 2014d, p. 4).

La evaluación institucional (p. 56) incluye la medición de la capacidad institucional de los gestores del PDA y de las entidades territoriales en cuanto a:

- Ejercer las funciones de planificación, regulación y supervisión del sector.
- Apoyar a los municipios en el fortalecimiento de capacidad empresarial y de gestión de las ESP.
- Desarrollar actividades de socialización del Programa y promover la participación de los ciudadanos en el desarrollo y supervisión de los proyectos.
- Flujo de las relaciones inter o intra institucionales que se dan en el Programa, y cómo inciden en el logro de los resultados.

Esto sumado a la medición de la capacidad institucional de los municipios que están vinculados a Bonos de Agua, o que no están vinculados ni al PDA ni al esquema de Bonos de Agua. Por su parte, la evaluación de resultados incluyó el abordaje de los siguientes puntos (DSEPP-DNP, 2014d, p. 57):

- Número de esquemas regionales constituidos.
- Número de operadores que manejan esquemas eficientes.
- Mejoramiento en las funciones de planificación, regulación y supervisión de los departamentos y gestores de PDA.
- Mejoramiento de los procesos de planeación de las inversiones.
- Mejoramiento de los indicadores de prestación del servicio.

Todos los componentes del estándar de objeto y alcance de la evaluación fueron adecuadamente direccionados en esta evaluación. Del mismo modo, y en consonancia con el alcance descrito, en el pliego de condiciones definitivo se describen los objetivos específicos de la evaluación (DSEPP-DNP, 2014d, p. 54), por lo que el estándar correspondiente a estos también fue satisfecho.

En cuanto a los componentes del estándar de oportunidad, pertinencia y utilización que pueden considerarse como reglas de alcance (los otros se verán en difusión, en reactividad/coordinación o en efectividad de la evaluación) se tiene que a pesar de que en el informe final no existe una clara diferenciación entre conclusiones, recomendaciones y lecciones aprendidas, en el PTIR si se realiza una clara diferenciación de estos puntos, y se presentaron de modo orientado a la utilización de la evaluación, por lo que estos componentes fueron satisfechos. De particular importancia para este punto fue la elaboración y presentación en el PTIR de la Matriz de recomendaciones (Economía Urbana e Infométrika, 2015b, p. 17), la cual relacionaba cada recomendación con las conclusiones de la evaluación, con las acciones de implementación y los actores responsables de estas.

4.2.2 Reglas de frontera y participantes

En las entrevistas realizadas (Asesora PDA) se pudo evidenciar que dentro del Comité de Diseño y el Comité de Seguimiento solo se incluyó la participación del MVCT como entidad sectorial implementadora del PDA[56]. La gran mayoría de las entidades encargadas de implementar las recomendaciones de la evaluación, no fueron incluidas en estos espacios. Estas organizaciones ausentes fueron: Gestores PDA, alcaldías, ESP, Superintendencia de Servicios Públicos, Comisión de Regulación de Agua Potable y Saneamiento Básico, y el Ministerio de Hacienda y Crédito Público. Esto implica que en esta evaluación no se satisficieron los estándares de participación de los actores relevantes, enfoque asociativo y de consideración de evaluación conjunta.

La modalidad de contratación de la evaluación fue concurso de méritos abierto. De ocho firmas consultoras que se presentaron, cinco cumplieron con los requisitos habilitantes, y de estas, la Unión Temporal entre Economía Urbana e Infometrika, fue la que resultó seleccionada con la más alta calificación en su propuesta de enfoque metodológico, en los demás criterios: recurso humano, experiencia y apoyo a la industria nacional, todos los postulantes habilitados obtuvieron la máxima calificación (Resolución 4287 de 2014 del DNP, pp. 5-6).

En el pliego de condiciones definitivo se aclaran la modalidad de selección (DSEPP-DNP, 2014d, p. 5), las personas que podrán participar, los criterios de calificación de las propuestas (pp. 22-35), y el equipo de trabajo mínimo (p. 16). Hubo un proceso de contratación abierto, se publicaron en el portal del SECOP las calificaciones a las firmas postulantes y hubo audiencia pública. El equipo requerido fue: el director de la evaluación; experto en el sector de política de acueducto y saneamiento básico; y el analista cuantitativo, cada uno con particulares requerimientos de nivel educativo y de experiencia específica. Todo esto satisfizo el estándar de equipo de la evaluación.

En este proceso de contratación existieron requisitos financieros habilitantes (índice de liquidez, índice de endeudamiento e índice de cobertura de intereses de la firma consultora), con estos se garantizó la capacidad de los proponentes de asumir compromisos y la posibilidad de cumplir con las obligaciones del contrato (DSEPP-DNP, 2014d, pp. 18-22). El valor del contrato fue de $1.142.430.554 para el DNP, subdivididos en 5 pagos. Con lo anterior, los componentes del estándar de recursos fueron completamente satisfechos.

[56] Y se confirmó con la respuesta a derecho de petición del DNP N° 20193600428271.

4.2.3 Reglas de elección y acciones

No se evidencia que existiera una reconstrucción exhaustiva de toda la teoría de la intervención previa a la realización de la evaluación, esto está estipulado en la normatividad general de la política de evaluación de políticas con la metodología de cadena de valor. Esto se refleja, particularmente, en algunos de los objetivos y en las preguntas de investigación que buscan satisfacer esta necesidad previa a la evaluación. Algunos ejemplos de esto son:

- "Realizar una revisión y análisis de los diferentes indicadores que existen en el sector para medir eficiencia y eficacia de la gestión administrativa, así como de la prestación del servicio, validando aquellos que permiten realizar una valoración más pertinente de los efectos del Programa" (DSEPP-DNP, 2014d, p. 54).
- "Realizar una revisión y análisis documental, con información existente, del Programa de Modernización Empresarial que permita evidenciar los principales cambios de los PDAs con respecto a este programa." (p. 54).
- "¿Cuáles son los indicadores más favorables para valorar la capacidad institucional del Programa?" (p. 54).
- "¿La base normativa y las políticas de desarrollo institucional del Programa establecen responsabilidades claras y estructuras e instrumentos que permitan hacer un el mejor uso de los recursos disponibles del sector?" (p. 55)
- "¿Cuál es la propuesta de indicadores y la metodología para implementarlos, que permiten una valoración más pertinente de los resultados del Programa?" (p. 55).
- "Debe describir el mapa institucional del Programa y la identificación de los roles de los actores del Programa" (p. 57).
- "Debe identificar la estructura general del Programa, definiendo con precisión su objetivo general; los resultados esperados con la intervención y las metas fijadas a cada resultado" (p. 57).
- "Debe identificar la estructura de funcionamiento del sector de agua y saneamiento básico en los municipios vinculados a Bonos de Agua y los municipios no vinculados a algún esquema" (p. 57).
- "Debe especificar con claridad las variables e indicadores o medidas observables a utilizar para comparar los resultados de los municipios vinculados a PDAs, los municipios vinculados a Bonos de Agua y los municipios no vinculados a ningún esquema" (p. 58).

Estos lineamientos evidencian que en el contrato de evaluación se estableció como responsabilidad del evaluador/contratista la definición de elementos fundamentales de la intervención, por lo que en la etapa de implementación de la evaluación se establecieron actividades correspondientes a la etapa de diseño de la evaluación, en este sentido, el estándar de evaluabilidad no fue satisfecho en este proceso de evaluación, puesto que se decidió emprender la evaluación sin tener una definición clara del objeto a evaluar en términos conceptuales y de operacionalización.

Al ser el PDA un programa que busca coordinar actores de varios sectores y de los tres niveles de gobierno alrededor de determinados procesos de planeación, la evaluación de este tuvo necesariamente que tener en cuenta la diversidad de planes, actividades y políticas nacionales y locales en el sector de agua y saneamiento básico. Un conocimiento profundo en estos aspectos es requerido puntualmente en el pliego de condiciones definitivo (p. 23). Esto satisfizo el estándar de coordinación y alineamiento.

En esta evaluación, la selección de la metodología no sucedió en el marco del proceso interno de diseño de la evaluación tal como está estipulado en la normatividad técnica y jurídica de la política. En este caso, el pliego de condiciones requirió al evaluador mayor protagonismo en todos los componentes metodológicos, en contraposición a únicamente proponer la metodología de recolección de datos (tal como manifestaron los funcionarios de la DSEPP). Dentro de los requerimientos puntuales de diseño de la evaluación se encontraron:

- "El CONSULTOR deberá proponer metodologías de evaluación según crea conveniente para el estudio, teniendo en cuenta que la metodología propuesta debe ser clara, precisa y suficientemente desarrollada frente a los objetivos de la evaluación, y en particular frente a la estrategia que se utilizará para generar las conclusiones y recomendaciones" (p. 55).
- "… debe tener en cuenta como mínimo (…) Claridad conceptual de los términos a emplear en la evaluación institucional" (DSEPP-DNP, 2014d, p. 56).
- "En términos de la justificación técnica de la metodología (…) Debe establecer la relación existente entre las preguntas orientadoras del componente, las variables de análisis de cada pregunta orientadora, los indicadores para cada variable, los instrumentos de recolección de información y los objetivos específicos de la evaluación" (p. 57).
- "Debe considerar la metodología más adecuada para identificar casos exitosos, casos promedio y casos no exitosos de PDAs" (p. 58).
- "Debe considerar la metodología más adecuada para identificar casos exitosos, casos promedio y casos no exitosos de municipios vinculados a Bonos de Agua y de municipios no vinculados a ningún esquema" (p. 58).
- "Debe definir claramente, luego de una revisión de indicadores del sector, las medidas observables a utilizar en cada variable de análisis" (p. 58).

A pesar de no cumplir a cabalidad la normatividad de la política de evaluación, y de no tener un diseño de evaluación previo a la contratación de esta, se observa que el propósito, alcance y preguntas evaluativas determinaron la metodología, hubo un informe inicial para explicar y justificar la selección del enfoque metodológico, y las técnicas de recolección y análisis de datos (p. 62), que se distingue entre niveles de resultados, y se realizaron requerimientos suficientes para la selección de indicadores. Dentro de estos últimos se incluyeron criterios de pertinencia,

sostenibilidad, eficiencia, eficacia (Economía Urbana e Infométrika, 2015a, p. 29). Con todo esto, se garantizó el cumplimiento de los estándares de selección de enfoque y método y de selección y aplicación de criterios de evaluación.

Las preguntas de investigación tienen coherencia con los objetivos de investigación y con la metodología seleccionada, por lo que el estándar de cuestiones a responder fue satisfecho. Entrando en su respuesta, dos de las preguntas que se encuentran en el pliego de condiciones no fueron respondidas en el informe final o en el PTIR, estas son:

- "¿Cuál es la diferencia de capacidad institucional entre municipios vinculados a PDAs y municipios vinculados a Bonos de Agua y entre municipios vinculados a PDAs y municipios no vinculados a ningún esquema?" (DSEPP-DNP, 2014d, p. 55).
- "¿Cuáles son las diferencias más relevantes en los resultados alcanzados por municipios vinculados a PDA y aquellos que se vincularon a Bonos de Agua y los que no se vincularon a ningún esquema, en términos de su capacidad para ejecutar los recursos, inversiones y cumplir con las metas de prestación de servicios?" (p. 55).

Por esta razón, solo puede considerarse como parcialmente satisfecho el estándar de respuestas a preguntas evaluativas. Adicional a esto, es de resaltar que muchas de las preguntas de evaluación no fueron contestadas ni se les realiza mención alguna en las conclusiones del informe final. Tampoco se presentan las recomendaciones de manera separada a las conclusiones o a los hallazgos, por lo que en este producto no se refleja la claridad del análisis realizado. Sin embargo, en el PTIR si se hace una clara diferenciación y se evidencia un flujo lógico entre hallazgos, conclusiones y recomendaciones, particularmente con la matriz de recomendaciones incluida en este producto (Economía Urbana e Infométrika, 2015b, p. 17). A pesar de no encontrarse propiamente en el informe final, el PTIR satisfizo el estándar de claridad del análisis.

El resumen ejecutivo de la evaluación se presentó como producto anexo al informe final, este presenta los principales hallazgos, pero no hace una clara diferenciación entre estos y las conclusiones, además no presenta las recomendaciones producto de la evaluación. Esto significa que no se satisfizo el estándar de claridad y representatividad del resumen.

Todos los componentes del estándar de contexto de la intervención, como el contexto de políticas, contexto de desarrollo, factores socioeconómicos, contexto institucional, y la participación de actores implicados, están abordados con las preguntas de investigación que corresponden a la evaluación institucional (DSEPP-DNP, 2014d, pp. 54-55). Adicional a esto, en el informe se aclara: "… se analizó la estructura institucional del Programa, su relación con el contexto y los resultados alcanzados teniendo en cuenta los procesos a nivel nacional y territorial, las interacciones entre los actores y los recursos que utiliza el Programa" (Economía Urbana e Infométrika, 2015a, p. 23).

En coherencia con las preguntas y los objetivos de la evaluación, en el informe final se hace una reconstrucción a profundidad de la lógica de la intervención (pp. 23-81). De particular importancia para la representación de la lógica de la intervención se encuentran dos esquemas: 1) fases de los PDA (p. 24), que incluye: diagnóstico, estructuración, e implementación y seguimiento; y 2) Delimitación sistémica de la política PDA (p. 46) que incluye: marco legal, insumos, intervinientes, y productos/servicios. Con lo anterior, el estándar de lógica de la intervención fue completamente satisfecho.

La metodología seleccionada y revisada por la supervisión del contrato se explica dentro del informe final de la evaluación (pp. 15-22), esto incluye una descripción detalla del muestreo, las fuentes de información primaria y secundaria, las técnicas y herramientas de recolección de datos (y a quienes van a ser aplicadas cada técnica), y las técnicas de análisis de datos, por lo que los estándares de explicación de la metodología y de validez y fiabilidad de las fuentes de información fueron satisfechos en todos sus componentes.

También se cumplió plenamente con el estándar de consulta y protección de implicados. Por un lado, se describe con suficiente detalle los actores de la política a los cuales se les debe aplicar los instrumentos de recolección de datos (DSEPP-DNP, 2014d, p. 58), lo cual permitió valorar la inclusión de todos los actores relevantes a ser consultados. Y por otro, se garantiza la protección de los derechos de estos actores al existir lineamientos claros de confidencialidad y de reserva y entrega de información (pp. 46-47).

No se realiza ninguna mención acerca de cambios en la metodología, y se realizan muy pocas frente a los límites de esta. Dentro de estos últimos se incluyen únicamente algunos comentarios acerca de la confiabilidad y del margen de error estadístico esperado según los grupos poblacionales bajo estudio (Economía Urbana e Infométrika, 2015a, p. 18). Sin embargo, en las entrevistas se reflejó evidencia acerca de obstáculos para la realización de un ejercicio libre y abierto (se presenta en reglas de agregación), lo cual nunca se menciona en el Informe final, con lo que el estándar de mención de cambio y límites puede considerarse como parcialmente satisfecho. No se realiza ninguna mención de discrepancias entre miembros del equipo evaluador, por lo que no se cumplió con este estándar.

En la cláusula quinta del contrato de consultoría se establece que el informe final de la evaluación debe ser entregado el 13 de agosto de 2015, y el PTIR el 7 de septiembre de 2015. Sin embargo, el informe final tiene fecha del 14 de septiembre y el PTIR del 28 de septiembre, por lo que hubo un retraso en la entrega de estos productos. En todo caso, por la naturaleza de las recomendaciones, así como del proceso de modificación del PDA, no puede considerarse que esto implicó un impacto negativo determinante en la relevancia temporal de los resultados y

recomendaciones de la evaluación. Por esto, a pesar del corto retraso, el estándar de evaluación dentro del plazo y el presupuesto previsto fue satisfecho.

4.2.4 Reglas de agregación y control

En la cláusula séptima del contrato de consultoría se establece como responsabilidad del director de la DSEPP, o de quien designe, la supervisión de las obligaciones contraídas por el consultor. Al igual que con los demás contratos, se establece que se conformará el Comité Técnico de seguimiento para el apoyo a la supervisión del contrato en lo que respecta a la supervisión de la calidad técnica de los productos del contrato. Así mismo que habrá una secretaría técnica dentro de este comité, la cual también estará a cargo de la DSEPP.

Adicionalmente, se establece que el supervisor debe aprobar los instrumentos de recolección de datos antes de que se inicie el trabajo de campo (DSEPP-DNP, 2014d, p. 62); que el consultor debe reportar cada quince días al supervisor del contrato, o a su delegado, el avance en el trabajo de campo en términos de recolección de información (p. 63); y en cuanto a la revisión de los productos y el pago al consultor:

> "Los informes de los cinco (5) productos descritos se entregarán en medio electrónico en las fechas establecidas en el cronograma de entrega de productos con base en la sección de "Plazo de ejecución" (…) De acuerdo con los comentarios y sus discusiones, el CONSULTOR realizará los ajustes solicitados por el Comité y enviará el informe ajustado en medio electrónico con el fin de que dicho Comité verifique los ajustes. Una vez aprobado, el CONSULTOR enviará el producto en físico y en medio magnético al supervisor del contrato para proceder con el recibido a satisfacción para pago" (p. 68).

Este contrato también incluyó reglas como el plazo máximo de una semana que tiene el consultor para entregar los productos con los comentarios que se le realicen como requisito para la expedición del recibo a satisfacción para su pago (DSEPP-DNP, 2014d, p. 4); los mecanismos de cobertura del riesgo (clausula décima segunda del contrato de consultoría); la cláusula penal compensatoria (décima quinta); y las multas (cláusula décima octava). Todo esto refleja una plena satisfacción de los estándares de control de calidad y de inclusión de observaciones de los implicados. Sin embargo, estas reglas, como se observó en el capítulo anterior, entran en conflicto con la independencia del evaluador.

4.2.5 Reglas de posición

En este proceso de evaluación también se observó limitada la comunicación entre los usuarios previstos y el evaluador/consultor. Esto porque el control que ejercieron los usuarios primarios (MVCT) estuvo restringido por la "consolidación, depuración

y tratamiento" que la DSEPP realiza a los comentarios que los miembros del Comité de Seguimiento hacen a los productos de la consultoría (DSEPP-DNP, 2014d, p. 52).

A pesar de que el evaluador contó con la posibilidad de defender (por una semana) sus comentarios ante el comité de seguimiento, las entrevistas reflejaron un grave problema:

> "En cuanto a la imparcialidad, conociendo que el director de la evaluación ya había estado en la formulación y en la implementación del programa, había muchas cosas que él presentaba y tenía el sesgo de: 'es que esto debieron haberlo hecho así'. Entonces nosotros le decíamos: 'si, pero no fue así y sabemos que no fue así'. La evaluación tuvo ese sesgo de: 'lo que pasa es que el programa no tendría tales falencias si ustedes lo hubieran ejecutado como se diseñó inicialmente'. Se hubiera hecho un poquito más imparcial, hubiera dicho: 'no mire, paso esto, después hubo estos cambios y se hizo así' hubiéramos tenido tal vez una conclusiones que no fueran tan sesgadas, que se ve un: 'no es que el programa decía que usted tenía que hacer eso, ¿por qué no se hizo?' y ¿por qué no se hizo? porque esto es una cosa de que los técnicos pueden decir que se haga de una forma pero aquí hay personas y también son asuntos políticos y uno tiene que lidiar con departamentos y administraciones municipales, así que en el desarrollo las cosas pueden cambiar. Así que en ese sentido si me parece que hubo un poco de sesgo en la evaluación" (Asesora PDA, MVCT).

Al observar los comentarios que se les realizaron a los productos y que quedaron consignados en los conceptos técnicos del informe final y del PTIR, se tiene que casi todos están direccionados a solicitar un mayor rigor al evaluador, y no a cambiar la esencia de sus valoraciones o recomendaciones. Frente a estas últimas se solicitó ampliarlas incluyendo las modificaciones institucionales específicas que se requieren para su implementación (DSEPP-DNP, 2015c, p. 3). A pesar de esto, debido a lo expresado por la funcionaria, no puede decirse que la evaluación tuviese una aproximación independiente de la gestión programática o de la formulación del PDA, lo que permite establecer que ni el estándar del proceso evaluativo libre y abierto, ni el de independencia del evaluador, fueron satisfechos en este proceso de evaluación.

4.2.6 Reglas de información

En el proyecto de pliego de condiciones, y el pliego de condiciones definitivo (posterior a la audiencia pública de adjudicación), como se ha visto incluyeron la finalidad, el alcance, los objetivos, un marco general para la selección de la metodología, los requisitos de presentación de informes, y fue consensuado entre

MVCT y el DNP, por lo que se cumplió el estándar de documentación acerca de la finalidad y la expectativa.

De la misma manera, se constata que el informe final de la evaluación cumple con los requisitos que se establecieron dentro del pliego de condiciones definitivo y dentro del contrato de consultoría, y a pesar de que no se organizan las conclusiones de tal manera que respondan a las preguntas de investigación, en general, se cumple con el estándar de informe de la evaluación, al ser este fácilmente comprensible.

En cuanto a la difusión que se le brindó a la evaluación, por un lado, se encuentran los lineamientos de la "Estrategia de socialización de insumos para la difusión de resultados y recomendaciones" del pliego de condiciones definitivo, en donde se establece la difusión de la evaluación a los miembros del Comité de Seguimiento, y la difusión que se realiza en el portal de SINERGIA en donde el informe está acompañado del resumen ejecutivo, y la ficha técnica de la evaluación.

Por otro lado, los esfuerzos en el cumplimiento de estos lineamientos de difusión de los resultados de la evaluación se revelan como limitados ante el elevado grado de rotación que ha existido en los últimos 4 años de los funcionarios encargados del diseño y coordinación del PDA al interior de MVCT. Frente a esto, los funcionarios de MVCT, manifiestan que, con el tiempo, muchos de los resultados se dejaron de socializar con las nuevas administraciones:

> "En este caso sí llegó [la información]. Sin embargo, también tenemos que tener en cuenta que este ministerio ha cambiado mucho de ministros durante los últimos dos años, entonces se ha hecho una socialización cada vez que ha llegado una persona de 'mire lo que pasa es que se hubo una evaluación y recomendaba esto y esto', y ellos han dicho 'si los vamos a tener en cuenta'. No obstante, si bien ellos pueden tenerlo en cuenta y estarlo revisando, en el momento en que haya un cambio ahí ya se pierde todo. Debido a eso no se volvió a hacer otra socialización" (Asesora PDA, MVCT).

En este sentido, existió y sigue existiendo una buena difusión externa de los resultados de la evaluación, pero el alcance de la difusión interna realizada en el marco del Comité de Seguimiento se vio fuertemente restringido por la rotación en la administración del PDA, con esto, solo puede considerarse como parcialmente satisfecho el estándar correspondiente a difusión de la evaluación.

4.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación del Programa Planes Departamentales de Agua

Uno de los grandes problemas de esta alta rotación en el personal encargado de la administración del PDA, fue que obstaculizó la utilidad que tiene el hallar e involucrar a los usuarios previstos desde etapas tempranas del proceso de evaluación. Al cambiar el personal administrativo y directivo, la subsecuente pérdida de memoria institucional conlleva una perdida en la apropiación que los usuarios previstos deben tener de la evaluación para garantizar su uso.

Tal como se planteó en las reglas de frontera, se contó con la participación del MVCT (quien es el líder de la intervención) desde las primeras etapas de la evaluación. Varias entidades con las cuales se deben implementar los PDA (Como la Superintendencia de Servicios Públicos Domiciliarios y las alcaldías) no participaron en los espacios de coordinación de esta evaluación. Sin embargo, esto no puede considerarse como una falencia en el hallazgo e involucramiento de los usuarios previstos. Por un lado, la participación de otros actores del PCI distintos al MVCT se dio de la siguiente manera:

> "Se hicieron reuniones también porque no fue solo el borrador [del nuevo decreto de los PDA)]. De hecho, se han hecho dos reuniones con gestores. Se hizo una reunión en el año 2016 en la cual se reunieron los gestores y les dijimos 'miren la evaluación dijo esto, ustedes que opinan'. Y nosotros también tomamos insumos de lo que ellos nos dijeron. Luego se hicieron otras reuniones con los gestores específicamente para unos temas de costos del gestor y de cómo iba a ser la participación del ministerio y esas cosas. Y después hicimos aquí al interior de nosotros otras reuniones. El año pasado se hizo otra reunión de gestores con la versión que ya teníamos del decreto y con ese decreto y los comentarios que ellos nos hicieron, se hizo esta otra versión y ahí si se circuló adentro. Pero nosotros lo hicimos en conjunto con ellos porque finalmente esto los iba a afectar y quien más que ellos que son los que están ejecutando la política" (Asesora PDA, MVCT).

El MVCT incluyó a los gestores PDA en la interpretación de los hallazgos de la evaluación, y como se observa, lo hace con el propósito de contar con insumos para la creación del nuevo decreto que vendría a regular el PDA. Adicionalmente, en el PTIR se observa que, de 17 recomendaciones, únicamente en una recomendación no se incluye como responsables al MVCT o a los gestores -que dependen directamente de este-, esta recomendación fue: "Reglamentar el artículo 6º de la ley 142 referido al cumplimiento de las condiciones para que proceda la prestación directa de servicios por parte de los municipios" (Economía Urbana e Infométrika, 2015b, p. 20). Tal como se verá más adelante, esta recomendación no fue implementada, lo cual es relacionada con la inclusión de los actores necesarios para

que fuera así. En resumen, se puede decir que si hubo un involucramiento satisfactorio de los usuarios previstos. Por otra parte, las entrevistas reflejan un balance positivo en cuanto a la dinámica del equipo de trabajo conjunto:

> “Pues yo creo que era una relación que funcionaba muy bien, porque eran personas que ya habían trabajado en otros temas juntos, entonces conocían de los planes departamentales y el sector, entonces digamos que era una relación armónica, en la que se planteaban las inquietudes que podía tener el ministerio y que podía tener el DNP, y se llegaba a acuerdos con el consorcio que realizó la evaluación” (Asesora PDA, MVCT).

Como se observa, la impresión que quedó por parte de los usuarios previstos es que el equipo de trabajo conjunto fue funcional. Una de las primeras tareas de este, fue definir los usos de la evaluación y las preguntas de investigación. En el mismo sentido, el balance alrededor de la reactividad del equipo evaluador y del DNP con respecto a las necesidades de los usuarios previstos fue positivo, puesto que se observa un interés por lograr acuerdos a partir de las necesidades que tienen estos últimos, y una fuerte coordinación entre el DNP y el MVCT:

> “El ministerio obviamente cuando veía cosas que no le parecía, o que veía que se podían arreglar, se planteaban las inquietudes y se realizaban. Pero quien marcaba la pauta de la información que se requería, que se quería obtener en la evaluación, era el DNP” (Asesora PDA, MVCT).

> “Se hicieron unas observaciones en una conversación previa que tuvimos ministerio y DNP, cuando nos presentaron [la selección de] la muestra, Entonces nos sentamos nosotros primero antes de ir allá, y nos dijeron '¿qué les parece?', y nosotros filtrábamos es esto y esto, y el DNP dijo 'listo' , y de ahí fue el DNP y dijo: 'tenemos esta posición', y el ministerio lo respaldó, pero hubo como una conversación previa nuestra para llegar a esta reunión de seguimiento [con el evaluador/consultor], digamos que así era como lo hacíamos, más que llegar al Comité y decir esto y esto, era como sentarnos y llegar a acuerdos con el DNP, y decir: 'venga necesitamos esto', llegamos a un acuerdo y con ese acuerdo se llegaba a la reunión con del Comité” (Asesora PDA, MVCT).

Como se observa, en este proceso de evaluación existió una buena disposición de la DSEPP de apoyar la inquietudes y preocupaciones de los usuarios previstos, y el equipo evaluador estuvo dispuesto a aceptarlas e incluirlas en los productos elaborados. Esta reactividad se extendió a posteriores etapas del proceso de evaluación. En cuanto a la adaptación del diseño de evaluación a la luz de las necesidades de los usuarios previstos, al preguntar acerca de transformaciones a la metodología presentada por el evaluador/consultor, se manifestó:

> “Sólo recuerdo una y era con respecto a la muestra. Esa es la única que recuerdo, y sé que se hizo, que fueron unas observaciones porque ellos no estaban de acuerdo con unas opiniones que nosotros teníamos con respecto

> a la muestra. Es fue la que yo recuerdo y que fue con la que hubo mayor discusión, y en la que tanto el DNP como el ministerio dijimos: 'no estamos de acuerdo por esto y esto', y si se tuvo en cuenta" (Asesora PDA, MVCT).

Lo único que restringe un balance completamente positivo en cuanto a la reactividad del evaluador/consultor o la DSEPP con respecto a los usuarios previstos, son los problemas de independencia ya mencionados:

> "La evaluación dio muchas herramientas y cosas que nosotros no teníamos presentes, es como cuando uno está tan metido en la ejecución de un programa no se da cuenta de ciertas cosas. Creo que un ojo externo pudo ver ciertas cosas que uno dice 'es verdad, eso está pasando'. Digamos que, si cumplió, aunque para mí no totalmente. A término personal, creo que la evaluación hubiera podido ser un poco más imparcial; pero si cumplió, si dio herramientas." (Asesora PDA, MVCT).

Esto sugiere poca independencia del evaluador, que surge de la relación laboral previa que tenía el director de la evaluación (equipo evaluador/contratista) con el MVCT para la formulación e implementación del PDA, como se observó más arriba (ver cita de entrevista al caracterizar estándar de independencia del evaluador). Si bien esta relación laboral previa facilitó la interacción dentro del equipo de trabajo conjunto, también implicó un sesgo que causó desacuerdos en el enfoque de la evaluación: "Digamos que uno esperaba más de eso. Yo creo que falto también conocer todo lo que se había desarrollado y tomarlo un poquito más en cuenta, más que decir que no se había ejecutado como se tenía que haber ejecutado" (Asesora PDA, MVCT).

En cuanto al involucramiento de los usuarios previstos en la construcción de recomendaciones e interpretación de los hallazgos, se debe entender que luego de la construcción de recomendaciones y conclusiones por parte del consultor evaluador, se realizó un panel de expertos que tuvo por objetivo la clasificación y priorización de las recomendaciones. Dentro de los participantes de este panel de expertos se incluyó a: expertos del sector del nivel nacional, entidades de regulación y vigilancia, académicos, exfuncionarios y consultores del sector (Economía Urbana e Infométrika, 2015b, p. 21). Los criterios de priorización fueron el nivel de incidencia que las recomendaciones tienen en el cumplimiento de los objetivos del programa, y la facilidad de implementación.

La metodología del panel de expertos fue a modo de taller con mesas de trabajo, en donde se presentó la escala de valoración con la cual se debían clasificar las recomendaciones, y se realizó un ejercicio de discusión en torno a las valoraciones de incidencia y facilidad de implementación. Luego: "… a partir de los resultados obtenidos en cada mesa, se realizó la clasificación y valoración definitiva de cada

una de las recomendaciones propuestas y se incluyeron las recomendaciones adicionales propuestas por los expertos en cada una de las mesas" (p. 14). Al realizar el cuestionamiento acerca de la participación de los usuarios previstos en este panel de expertos, se comentó:

> "Al panel asistimos, asistió el DNP y asistieron otros actores, sin embargo, no creo que hubiéramos visto mucho esas recomendaciones. Una de las cosas que habían dicho era esta conclusión de acabar el programa, y por ejemplo, de eso no recibimos como una ampliación de esto. (...) recuerdo el panel, la discusión, los planteamientos que tenía cada parte del programa, pero no recuerdo bien que se hubiera tenido en cuenta nuestras recomendaciones, creo que una de las conclusiones fue esa, necesitamos unas soluciones más concretas, eso se les dijo y no recibimos una respuesta a esto" (Asesora PDA, MVCT).

> "… creo que teníamos tanto DNP como ministerio, una mayor expectativa de las propuestas, de las conclusiones y recomendaciones (…) esperábamos como una propuesta más concreta de qué medidas se podían tomar (…) Digamos que las recomendaciones fueron buenas pero ¿qué se esperó en algún momento? que había algunas cosas que nosotros esperábamos que ellos desarrollaran, que no tuviéramos que desarrollar nosotros, pero si es verdad que nosotros si tomamos muchas de las cosas, que dijimos 'esto no lo habíamos visto" (Asesora PDA, MVCT).

Esto evidencia que, en este caso, los usuarios previstos no tuvieron un mayor interés en ser involucrados en la construcción de recomendaciones, puesto que se esperaba que los evaluadores se encargaran enteramente de este asunto. Aun así, su participación en el panel de expertos y los comentarios que se debieron hacer a los últimos productos, ciertamente garantizaron su involucramiento. En cuanto a estos últimos, se incluyeron principalmente comentarios orientados a resaltar el papel del índice de capacidad institucional, inclusión de temas, fortalecimiento de explicaciones, explicitaciones de opiniones de actores, homogenización del análisis entre casos (DSEPP-DNP, 2015d, pp. 4-5) para el producto del informe final. Y comentarios enfocados al desarrollo de los mecanismos de implementación de las recomendaciones, aspectos institucionales a desarrollar, y explicitación de aspectos que ya están en la política pero que aún no se han implementado (2015c, p. 3) para el PTIR. Todos fueron incluidos en los productos.

En este panel de expertos, se discute el costo, los beneficios y los desafíos de implementar las recomendaciones (en lo cual se halla la utilidad teórica de involucrar a los usuarios previstos en la construcción de recomendaciones). Al preguntarse por un posible vínculo con la DSEPP luego de la evaluación, se respondió que no hubo tal contacto, por esta razón, no hubo plan conjunto de seguimiento a la implementación de los hallazgos de la evaluación, ni búsqueda proactiva del uso por parte de los miembros del equipo evaluador.

4.4 Análisis de la efectividad de la evaluación del Programa Planes Departamentales de Agua.

En total, a partir de las conclusiones de esta evaluación se formularon 17 recomendaciones. De estas, 11 fueron tomadas en cuenta por parte de MVCT para la elaboración del nuevo proyecto de decreto con el que se buscaría implementar esas recomendaciones. De estas recomendaciones utilizadas, tres pueden ser agrupadas como cambios en la institucionalidad de los gestores PDA: “Definir con mayor precisión el alcance del rol que desempeña el Gestor como ejecutor de los PDAs” (Economía Urbana e Infométrika, 2015b, p. 6); “Definir el alcance del rol que desempeña el Gestor en temas ambientales” (p. 6); “Definir estructuras tipo para los gestores” (p. 6). Al respecto, la asesora líder de la formulación del PDA comenta:

> “En lo principal que se ha usado [la evaluación] es en el proyecto que se está desarrollando en este momento de la modificación del decreto de los PDA. Dentro de ese decreto se tomaron cosas relacionadas con el rol del gestor y la parte de estructuras tipo. No se ha hace una estructura tipo muy rígida sino se le están dando como unos lineamientos y unas necesidades que nosotros vemos para que ellos cumplan con sus funciones. La evaluación muestra que muchos no tienen los profesionales para cada una de las cosas que se desarrollan, y un solo profesional sin carrera de varios temas. Entonces, la idea era darle una línea para que por lo menos hubiera responsables y no se ponía una dependencia sin un profesional que estuviera al frente de estos temas. Eso es lo que se está buscando” (Asesora PDA, MVCT).

En este sentido, y teniendo en consideración la alta probabilidad de que este decreto sea implementado con estos elementos (lo cual es confirmado por la entrevistada), estas recomendaciones fueron implementadas (Por lo menos por el nivel técnico hasta 2019). Otra recomendación que se tomó en consideración para la construcción del proyecto de decreto fue promover contratos temporales con operadores especializados (de servicios públicos) para brindar asistencia técnica y operativa a los prestadores de servicios públicos que lo requieran (Economía Urbana e Infométrika, 2015b, p. 7). Frente a esto se responde: “En cuanto a contratos temporales con operadores especializados en asistencia técnica lo que se ve es que el gestor, muchos de ellos, no todos, están adquiriendo una experiencia y tienen un conocimiento y pueden hacer la parte de asistencia técnica. Eso es una cosa que también se está desarrollando en el marco del decreto” (Asesora PDA, MVCT).

La siguiente recomendación que se utilizó, corresponde a fortalecer los procesos ejecutados por el mecanismo de viabilización de proyectos, en donde se sugiere estandarizar los criterios de viabilización que elimine la subjetividad de los evaluadores de los proyectos PDA; priorizar la evaluación de proyectos con fuentes de financiación definida; y fortalecer la capacidad operativa de la ventanilla única de

MVCT (Economía Urbana e Infométrika, 2015b, p. 10). En las entrevistas se aclara el modo particular en que esto fue implementado:

> "Eso se ha venido haciendo y se han hecho planes de choque para agilizar todo este proceso de viabilización. La otra es que ahí [la recomendación] no tiene la viabilización en el mecanismo departamental, que hace que ya no lleguen todos los proyectos sino que ellos allá creen sus... pues le decimos ventanillas porque inicialmente eso se llamaba ventanilla única, pero creen mecanismos de viabilización, y ellos allá expidan los conceptos cuando tienen recursos de ellos. En este momento muchos tienen más recursos de ellos que de la Nación, por ende, no requieren venir acá. No lo hicimos mejorando aquí este mecanismo del ministerio, pero sí dándoles a ellos la posibilidad de que aprobaran sus proyectos alla y fuera un proceso mucho más rápido. Porque además esta recomendación está relacionada con las quejas de los municipios y de los gestores sobre la demora del ministerio para viabilizar. Aquí se ha intentado siempre agilizar el proceso debido a la necesidad, pero es que los proyectos que llegan acá son muchos, y evaluar todo eso requiere de muchas personas y de mucho tiempo. Lo que se ha hecho es priorizar para sacar más rápido lo que es más urgente; pero la ventanilla departamental ha servido para reducir esos tiempos y que ellos aprueben los proyectos en el departamento. Creo que en ese sentido eso fue una forma de implementarla" (Asesora PDA, MVCT).

Otras recomendaciones que se pueden agrupar y que reflejan el uso de la evaluación, corresponden a articular la ejecución de las inversiones a nivel regional, a partir de supeditar recursos de la nación al compromiso de los recursos por parte de los municipios (Economía Urbana e Infométrika, 2015b, p. 28), en conjunto con apoyar la planeación de largo plazo y articulada entre nación, departamentos y municipios (p. 7). Al respecto, se comentó:

> "…la planeación está muy relacionada con la ejecución de las inversiones (…) si nosotros miramos, la forma como funcionan los PDAs busca eso, que sea el gestor, o que sea el municipio [que haga la planeación], porque ahí hay una opción de que los municipios puedan ejecutar. Entonces creo que esta parte también está relacionada con las inversiones que se hacen a través de fondo de adaptación, de DPS y de otras entidades, yo creo que es más en ese sentido. Esa parte se quiere hacer también con lo de la planeación y lo que yo te decía arriba, que sea el gestor el que planee, el que haga todas estas inversiones. De hecho, lo que se ha dicho en los últimos comités en los departamentos es 'usted tiene proyectos de regalías ¿por qué no usa su gestor para que sea su estructurador de proyectos? que más que el gestor que es el que conoce cómo está la situación del sector en su departamento y en los municipios, para que el estructure y haga todos esos proyectos'. Se ha hecho la recomendación y hay varios casos en los que es el departamento o el gestor quien estructura los proyectos de regalías (…) Creo que con eso hemos logrado tomar un poco esa recomendación de inversiones a nivel regional, de su articulación" (Asesora PDA, MVCT).

La octava recomendación implementada sugiere fortalecer las actividades de fortalecimiento institucional realizadas por el gestor PDA, y que la cofinanciación de las obras que realice la nación esté supeditada a las recomendaciones del gestor PDA en cuanto a fortalecimiento institucional (Economía Urbana e Infométrika, 2015b, p. 10). En cuanto a esto, la asesora líder de formulación del PDA aclara:

> "Con respecto a eso también en el decreto se está intentando modificar un poco el plan de aseguramiento para que el gestor, según sus funciones, de un poco más de apoyo a los municipios. Hay unos planes de aseguramiento, el gestor los ejecuta y tiene unas actividades de fortalecimiento; algunas son continuas. Entonces lo que nosotros pusimos en las nuevas funciones, en las funciones ajustadas del gestor en el nuevo decreto que esperamos salga pronto, son de ayudarlos en cosas más pequeñas, como un apoyo técnico cuando ellos lo requieran. Es como prestar asistencia técnica y lo tenemos ahí como general para que den el apoyo a todas las necesidades; por ejemplo, si requieren asistencia técnica financiera o de cualquier tipo. Sin embargo, también hay que tener en cuenta que hay unos que no lo van a poder hacer de la misma forma que otros pues algunos tienen una mayor capacidad institucional. Es una cosa que hay que revisar, y a los que son más débiles tocaría hacerles como una capacitación para que puedan apoyar, o buscar la forma en que nosotros podamos apoyarlos a ellos. Eso es una cosa que está en construcción" (Asesora PDA, MVCT).

La novena recomendación utilizada correspondió a establecer los alcances de los actores en las actividades de asistencia técnica (Economía Urbana e Infométrika, 2015b, p. 10). La funcionaría líder de formulación manifestó que esto se está tomando en cuenta para el proyecto de decreto, en donde se le establecen funciones de asistencia técnica del MVCT a los gestores, y de los gestores a los municipios.

En décimo lugar, recomendó redefinir el rol del gestor frente al proceso de certificación de los municipios (p. 10), en cuanto a este, se manifiesta que si bien se tuvo en cuenta para la construcción del proyecto de decreto, no han profundizado en esto por el interés que existe en el sector de eliminar la certificación de los municipios: "… esa parte se conserva en este momento en el decreto y está, pero pues no le ponemos tanta atención porque sabemos que la certificación se quiere eliminar, que la idea del sector es eso, entonces por eso no le hemos prestado atención" (Asesora PDA, MVCT).

Por último, dentro de las recomendaciones que reflejan la utilización de la evaluación se encuentra: "Implementar mecanismos de seguimiento periódico a los instrumentos de planeación" (Economía Urbana e Infométrika, 2015b, p. 9). En este caso se respondió: "En el decreto ya se incluyó como tal presentar informes de todos los sistemas de planeación, antes en el decreto sólo estaba del PAEI, ahora si los

informes de seguimiento si van a quedar para cada uno de los sistemas de planeación. Si se venían haciendo cosas [antes de la evaluación], pero no estaban en el decreto" (Asesora PDA, MVCT).

Si bien se puede decir que los usuarios previstos identificados por el proceso de evaluación utilizaron las conclusiones y recomendaciones relacionadas con los anteriores temas, principalmente para la elaboración del nuevo decreto del PDA, es importante aclarar que a marzo de 2019 (cuatro años y medio después de la finalización de la evaluación), al preguntar por este decreto se respondió: "No ha salido (…) Aquí [MVCT] ya surtió el trámite, el proceso está ahora en planeación nacional, porque como es un proceso conjunto, digamos que todavía estamos en el trámite, no es que se haya caído" (Asesora PAD, MVCT). En este sentido, las recomendaciones fueron utilizadas y aplicadas por los usuarios previstos, a pesar de que aún no estén completamente implementadas.

En contraste con las recomendaciones que han sido (por lo menos parcialmente) implementadas, se encuentran seis recomendaciones que no fueron utilizadas. Por una parte, se recomendaron acciones que no pueden ser implementadas por estar dirigidas a esferas políticas, y por ende, quedaron por fuera del alcance del MVCT para poder implementarlas, estas son: "Fortalecer la institucionalidad enmarcada en la Ley 142 de 1994 para apoyar a los municipios y prestadores, e intervenir donde sea pertinente para asegurar la prestación eficiente de los servicios" (Economía Urbana e Infométrika, 2015b, p. 4), y " Reglamentar el artículo 6º de la ley 142 referido al cumplimiento de las condiciones para que proceda la prestación directa de servicios por parte de los municipios" (p. 4). Frente a estas, se manifestó:

> "… [El PTIR] no decía cómo implementarlas. Por ejemplo, la que decía reglamentar el artículo 6, ¿a qué se refiere con reglamentarlo? si el artículo ya está y dice "usted lo que tiene que hacer es tal y tal paso", ¿a qué se refería? Hubo muchas cosas que fueron muy generales. Eso es como la de terminar el programa, ¿con que lo voy a reemplazar? ¿qué voy a hacer? Había unas que eran muy generales y no decían como un paso a paso qué hacer, o cuál era la recomendación en términos puntuales" (Asesora PDA, MVCT).

> "…por ejemplo todo lo relacionado con la Ley 142 y su reglamentación. (…) esta ley necesita una actualización que se adapte a todo esto que ha ido pasando durante todos estos años y a las necesidades actuales del servicio. Eso son cosas que uno dice 'si', reglamentar por ejemplo el artículo 6, sí. Pero yo creo que más que eso, va a llegar un punto en que nos va tocar es modificar la misma ley (…) pero se requiere el respaldo político para poder hacer ese tipo de modificaciones. Entonces uno puede iniciar todo eso y decir que toca hacerlo, pero también toca que alguien del nivel superior tome la decisión de meterse en eso y modificarla" (Asesora PDA, MVCT).

Otra recomendación que tuvo inconvenientes para ser implementada corresponde a "Formular un mapeo de criterios para la distribución de los recursos de la Nación

y los departamentos en los municipios vinculados al PAP-PDA de manera articulada (Nación-Departamentos-Municipio)" (Economía Urbana e Infométrika, 2015b, p. 8). El tema de distribución de recursos de la nación, además de quedar por fuera del alcance del MVCT, desconoce los intereses que tiene el sector de agua y saneamiento básico:

> "… meternos con criterios de distribución solamente para los planes departamentales, pues sería una cosa limitada. Si nosotros vamos a hacer un criterio de distribución se tiene que hacer para todo el país, sin importar el programa o la política. Eso es algo que como sector se sabe que se necesita, pero requiere mucho desarrollo. Además, porque no lo podemos tomar solo para los planes departamentales. Yo creo que es algo que sí está en la idea del sector, pero no se ha desarrollado porque la idea es desarrollarlo para todo el sector, no solamente para el programa" (Asesora PDA, MVCT).

Una situación similar se presenta con la recomendación que plantea: "Ejecutar recursos de la Nación vía PDAs y condicionarla al compromiso de recursos de los municipios" (Economía Urbana e Infométrika, 2015b, p. 8). En este caso se manifestó: "… pero eso ya es como ministro: 'señor presidente cómo va a repartir usted la plata' Eso no es como tan abajo nuestro (…) es una decisión política (…) creo que la evaluación tiene bastantes recomendaciones políticas. Tiene varias" (Asesora PDA, MVCT). También sucede con la recomendación: "Instituir un mecanismo de acceso a crédito para todos los municipios" (p. 8), la cual

> "… depende mucho del nivel superior y la voluntad política. Cuando nacieron los planes departamentales estaba el apalancamiento a través del FIA con las vigencias futuras, cuando la Contraloría dice 'eso no se puede hacer', digamos que ahí se tranca todo ese endeudamiento. Entonces (…) necesitamos de otros actores que ayuden a construir todo esto" (Asesora PDA, MVCT).

Por último, se recomendó promover la sostenibilidad financiara de los prestadores de servicios públicos, esto a través de destinar recursos para cubrir sus costos operacionales de manera temporal (p. 11), en este caso se manifestó lo siguiente:

> "… eso se puede promover a través de los planes de aseguramiento de fortalecimiento institucional a la transformación empresarial. Porque uno lo que busca es mirar cuáles son las falencias o las dificultades que tienen las empresas para ser sostenibles. Esto se está haciendo, ahí dice promover pero eso se viene haciendo desde siempre porque lo que buscamos es eso. Los planes departamentales son planes para el manejo empresarial de los servicios. Esos planes de aseguramiento buscan es eso (…) es algo que se viene haciendo, siempre se ha hecho. Lo que pasa es que no es sencillo. Además porque tú tienes que contar con la voluntad política del Alcalde. Tú

> puedes decir ‘Alcalde la mejor opción es que haga esto’; pero si él no quiere adoptarlo o el consejo no le aprueba eso, no va a salir”.

Como ya existían acciones dirigidas a promover la sostenibilidad financiera de las ESP, esta recomendación resulta irrelevante. En síntesis, a pesar de la formulación de muchas recomendaciones que no pudieron ser implementadas por estar dirigidas a otros actores que no fueron incluidos en el proceso de evaluación, la gran mayoría de las recomendaciones que estuvieron dirigidas a MVCT fueron utilizadas, por lo que en general, el uso de esta evaluación ha sido positivo.

4.5 Conclusiones

En cuanto al análisis de aplicación de estándares de calidad, se encuentran problemas en los siguientes estándares. Dentro de las reglas de frontera, no se satisficieron los estándares de participación de los actores relevantes, enfoque asociativo y de consideración de evaluación conjunta, al dejar por fuera del proceso evaluativo a gran cantidad de actores con los cuales se coordina MVCT para la formulación e implementación del PDA.

Frente a reglas de elección, no se satisfizo el estándar de evaluabilidad al decidir emprender la evaluación sin tener una definición clara del objeto a evaluar en términos conceptuales y de operacionalización; solo puede considerarse como parcialmente satisfecho el estándar de respuestas a preguntas evaluativas por la desconexión que existe entre las preguntas y las conclusiones del estudio; hubo muy pocas menciones de cambios y límites de la metodología de la evaluación, y no se realizó ninguna mención de discrepancias entre miembros del equipo evaluador, por lo que no se cumplió con este estándar.

Con respecto a los estándares que corresponden a reglas de posición, el elevado grado de control de la DSEPP sobre los productos del evaluador/consultor, si bien garantizar un adecuado control de calidad, llegan al punto de ser un obstáculo para los estándares de independencia del evaluador, y de proceso evaluativo libre y abierto. Este problema se incrementó al contar con un director de la evaluación (dentro del equipo del evaluador/consultor) que en el pasado había trabajado en la formulación e implementación del PDA dentro de MVCT. Por último, en las reglas de información, se reflejaron problemas dentro del estándar de difusión, no por problemas dentro del diseño de la difusión o por un escaso esfuerzo en esta labor por parte de los usuarios primarios o del DNP, sino por el elevado grado de rotación en el personal directivo de MVCT, lo cual afectó negativamente el grado de apropiación necesario para la evaluación.

En relación con los patrones de coordinación necesarios para que el proceso de evaluación sea efectivo, se encontró que se tuvo cuidado, al momento de formular las recomendaciones, de conservar dentro del rol de usuario previsto al MVCT. Se evidenció un buen desenvolvimiento del equipo de trabajo conjunto, una relación

armónica, en donde se planteaban inquietudes y se llegaba a acuerdos, y en general, una buena reactividad de la DSEPP y del evaluador/consultor frente a las inquietudes y necesidades de los usuarios previstos. Esto se vio reflejado en las discusiones metodológicas y en la adaptación de la metodología según las necesidades de los usuarios. También se evidencia un involucramiento de los usuarios previstos en la interpretación de los hallazgos y en la construcción de las recomendaciones. No existió una planeación conjunta de seguimiento a la implementación de los hallazgos, ni una búsqueda proactiva del uso.

A pesar de los problemas dentro de los estándares de calidad y de coordinación, el balance general del uso de la evaluación, y por ende, en la efectividad del proceso evaluativo puede describirse como positivo. 11 de 17 recomendaciones han sido (parcialmente) implementadas por parte de MVCT, y los gestores PDA. Es importante resaltar que las recomendaciones que no fueron implementadas tuvieron el problema común de estar dirigidas a instancias de gobierno distintas a los usuarios previstos con los que se trabajó a lo largo de la evaluación, por lo que pueden ser descritas como irrelevantes.

Esto resalta la importancia de los estándares correspondientes a reglas de frontera que no fueron satisfechos en este proceso de evaluación, al hallazgo e involucramiento en todo el proceso de evaluación de usuarios previstos con poder e influencia y con suficiente interés, y al involucramiento de los usuarios en la construcción de recomendaciones. Estos últimos, como componentes de la coordinación entre el proceso de evaluación con el de política pública.

Esta evaluación se realizó en un contexto con fuerte interés en la reformulación del programa a través de la formulación de un decreto modificatorio al Decreto 2246 de 2012. A julio de 2019 se cuenta con proyecto de decreto en proceso de sanción presidencial. Como se observó, muchas de las recomendaciones de la evaluación fueron implementadas a través de ese nuevo decreto, por lo que la evaluación estuvo disponible en un escenario en donde fue considerada como necesaria su información. Esto significó que hubiera un alto grado de apertura hacia el cambio de la política con base en la información de la evaluación por parte de los usuarios previstos.

5. Efectividad de la Evaluación de Operaciones de la Política de Atención Humanitaria a la Población en Situación de Desplazamiento

5.1 Presentación de la Política de Atención Humanitaria a la Población en Situación de Desplazamiento

Actualmente la Política de Atención Humanitaria es un componente dentro de toda la Política de Atención y Reparación Integral a las Víctimas del Conflicto Armado cuya institucionalidad principal se encuentra en la Ley 1448 de 2011. En un principio, la Política de Atención Humanitaria surge a partir de la Ley 337 de 1997 la cual establece (art. 15) como regla de alcance de la política el "… garantizar la atención humanitaria de emergencia con la finalidad de socorrer, asistir y proteger a la población desplazada y atender sus necesidades de alimentación, aseo personal, manejo de abastecimientos, utensilios de cocina, atención médica y psicológica, transporte de emergencia y alojamiento transitorio en condiciones dignas".

Un hito importante de esta política, así como de todas las que están dirigidas a la atención y reparación a víctimas, fue la Sentencia de tutela T025 de 2004 de la Corte Constitucional. En esta sentencia se declaró un estado de cosas inconstitucional, principalmente por la vulneración masiva y generalizada de derechos constitucionales; por la prolongada omisión de las autoridades en el cumplimiento de sus obligaciones; la no expedición de medidas legislativas, administrativas o presupuestales para evitar estas vulneraciones; y la existencia de un problema que requiere la intervención coordinada de múltiples entidades en diversos niveles de gobierno.

En esta sentencia y con respecto a la atención humanitaria, la Corte Constitucional estableció que el derecho a la subsistencia mínima es una expresión del derecho al mínimo vital, el cual es un deber del Estado y que debe incluir: alimentos esenciales y agua potable; alojamiento, vestido apropiado, servicios médicos y sanitarios. A partir de la Ley 1448, esta atención humanitaria se dividirá en tres clases, las cuales se deben asignar dependiendo de la situación particular de la persona víctima:

- Atención inmediata: la cual se entrega a personas u hogares en situación de desplazamiento y que se encuentran en situación de vulnerabilidad acentuada y requieren alberge temporal y asistencia alimentaria. Esta atención es entregada por parte de los entes territoriales (Econometría y Sistemas Especializados de Información –SEI-, 2014a, p. 29).
- Atención de emergencia: es entregada a personas y hogares en situación de desplazamiento una vez se expide el acto administrativo que los incluye en el Registro Único de Víctimas (RUV). Esta se entrega dependiendo de la necesidad y urgencia con respecto a la subsistencia mínima, está a cargo de la Unidad Administrativa Especial para la Atención y Reparación Integral a las Víctimas –UARIV- (p. 29).
- Atención humanitaria de transición: Esta es entregada a personas en situación de desplazamiento que están incluidos en el RUV, que no cuentan con elementos necesarios para su subsistencia mínima, pero cuya situación no presenta gravedad y urgencia. Está a cargo de la UARIV, el Instituto Colombiano de Bienestar Familiar (ICBF), en conjunto con los entes territoriales (p. 30).

Dentro de su reglamentación se encuentra la Resolución 00171 de 2014 acerca de procesos de entrega y reconocimiento de atención humanitaria; Resolución 03069 de 2010 acerca de la entrega de atención humanitaria a la población; circular 001 de 2010 sobre los valores componentes de atención humanitaria; resolución 1956 de 2012 sobre el procedimiento para la solicitud y el trámite para la entrega de la atención humanitaria; Decreto 4800 de 2011 que reglamenta la Ley 1448 de 2011; y el Decreto 2569 de 2014.

5.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Atención Humanitaria

5.2.1 Reglas de alcance y resultados potenciales

La UARIV y el DNP identificaron que a pesar de que se brinda los componentes de la atención humanitaria a las personas en situación de desplazamiento, no es posible afirmar que gocen del derecho a la subsistencia mínima (DNP, 2014c, p. 39). Esto se ha reflejado en la continua solicitud de Atención Humanitaria de Transición por parte de la población desplazadas, quienes se limitan a acudir a esta sin acceder a los otros mecanismos de estabilización socioeconómica que les son destinados.

Ante esta situación, se consideró necesario indagar acerca del grado de satisfacción de los requerimientos básicos de la subsistencia mínima que se obtiene a partir de la entrega monetaria; la relación entre el estado de articulación institucional y la

oportunidad, suficiencia e integralidad de la atención; la pertinencia o necesidad de este mecanismo; y sus posibles complementos e incentivos (Econometría y SEI, 2014a, p. 3). A esta demanda busca responder la evaluación de operaciones.

Se manifestó -como requerimiento- recomendaciones con respecto a la superación de la fase de transición a partir del diseño de un mecanismo de suspensión de la Atención Humanitaria de Transición, que permita verificar y caracterizar las verdaderas necesidades en materia de atención o asistencia de los beneficiarios (DSEPP-DNP, 2012c, p. 3). Esto caracteriza el uso que se pensó para esta evaluación, y el tipo de desarrollo de capacidad que se esperó obtener de esta.

Es importante aclarar que no se esperó un diseño de este mecanismo de suspensión por parte del equipo evaluador. En contraste, las recomendaciones que se solicitaron dentro del diseño de la evaluación se direccionaron a que ayudaran a validar el nuevo diseño de atención humanitaria con el que la UARIV contaba al momento de realizar la evaluación:

> "En la actualidad la Unidad para la Atención y Reparación Integral a las Víctimas, en particular la Dirección de Gestión Social y Humanitaria, cuenta con el diseño de la nueva política de Atención Humanitaria. Por lo tanto, esta evaluación es fundamental para validar el diseño de la política en términos de las definiciones técnicas utilizadas, el proceso integral de la solicitud y entrega, los efectos e incentivos que ha generado su implementación, y finalmente para conocer el estado del goce efectivo del derecho a la subsistencia mínima de los hogares víctimas de desplazamiento. Además de esto, se podrán establecer cuáles han sido los vacíos de la política e identificar en qué medida los lineamientos de la nueva política solucionan los problemas encontrados en la valoración de la operación del modelo actual." (Banco Mundial, 2014, Pág. 1)

Como se observa, esta evaluación consistió en una mezcla entre una evaluación ex-ante del nuevo modelo y ex-post del modelo de atención humanitaria que se utilizó hasta 2014. Se manifestó puntualmente: "[DNP] requiere contratar una firma consultora que desarrolle la evaluación (…) cuyo fin último es proponer recomendaciones para el diseño del nuevo modelo. (DNP, 2014c, p. 40).

La combinación entre ambos tipos de evaluación se realiza a partir de indagar acerca de la pertinencia del nuevo modelo para dar respuesta a los problemas del modelo que se tenía hasta la fecha, y si de esta manera permitía "… garantizar el derecho a la subsistencia mínima de la población desplazada en el momento y de la forma adecuada" (p. 39). Más adelante se evidencia con los datos adicionales, el grado en que esto significó o no, una justificación o legitimación de decisiones previas. Hasta este punto, sólo se puede decir que esto satisfizo los estándares de fundamentación y finalidad de la evaluación y de desarrollo de capacidad.

En el contrato de la consultoría se establece como objeto general: "Realizar una evaluación de operaciones a partir de la valoración de los procesos, productos y

resultados del modelo actual de la Política de Atención Humanitaria a población desplazada con el fin de analizar la pertinencia del diseño de la nueva política." (p. 40). Junto con esto, se realiza una descripción de la política de atención humanitaria, con todos sus componentes (p. 37), los plazos y condiciones de pago (p. 30), la zona geográfica, y las demás dimensiones a abordar por parte de la evaluación.

Particularmente, la evaluación estuvo estructurada a partir de cuatro componentes. El primero se dirigió a abordar la estructura operativa del modelo actual de la política, cuya finalidad fue analizar: "… la operación de la política de Atención Humanitaria Inmediata, de Emergencia y de Transición y la identificación de las fortalezas y de los vacíos en la prestación del servicio a la Población Desplazada (PD), tanto a nivel nacional como territorial." (p. 40). En este componente se aborda la valoración del proceso de atención humanitaria con respecto a sus propósitos, la identificación de cuellos de botella en el proceso, la percepción de funcionarios y beneficiarios, los incentivos que se generan, y la caracterización de los hogares beneficiarios.

En segundo lugar, se estableció el componente acerca del nuevo modelo. En este se buscó abordar la pertinencia del diseño del nuevo modelo a partir de: "…la valoración de los procesos, productos y resultados de la política actual. Es así, como se debe identificar si los lineamientos que fueron diseñados para la nueva política solucionan los problemas encontrados en la valoración de la operación del modelo de atención actual" (p. 43).

El tercer componente de la evaluación se dirige a establecer las reglas a partir de las cuales se debieron construir las conclusiones y recomendaciones de la evaluación. Su objetivo fue analizar la información y resultados obtenidos de los anteriores componentes, para identificar los aspectos de la operación de la política que deben ser ajustados (p. 44). Establece que las conclusiones y recomendaciones deben estar soportadas tanto por análisis cualitativos como cuantitativos, y que estas ayuden a ajustar el nuevo diseño con el que se cuenta en ese momento.

Por último, en el cuarto componente se especifican reglas adicionales para el PTIR. Particularmente se aborda el diseño de la estrategia de socialización y uso de los resultados de la consultoría, el cual "…le permita a la Unidad para las Víctimas utilizar dicha información en la toma de decisiones y ajuste de la Política. Por tanto, el consultor deberá especificar la forma más apropiada para realizar la transferencia de conocimiento y específicamente los usos del nuevo diseño ajustado de la Política" (p.44). Esta información revela la satisfacción de los estándares de objeto y alcance de la evaluación y de objetivos específicos.

Se debe resaltar que las recomendaciones que se requieren en el contrato de consultoría se dirigen a ajustar el diseño del nuevo modelo de atención humanitaria (p. 45), por lo que, si bien se establece dentro de la finalidad de la evaluación, la

validación del nuevo diseño, se direccionan las recomendaciones a no únicamente validar, sino “ajustar” o perfeccionar el nuevo diseño. Esto es importante porque en la medida en que las recomendaciones apunten a este objetivo, y se implementen se estaría racionalizando el proceso de toma de decisiones y no únicamente justificando decisiones previas con la evaluación.

Dentro de los lineamientos de estas recomendaciones, en el contrato se establece que estas debían ser analizadas en el marco del Comité de Seguimiento, y posteriormente, presentadas a los directivos de las respectivas dependencias del DNP con el fin de obtener incidencia, analizar su oportunidad y precisar la forma de su incorporación (p. 59). También, que estas deben ser presentadas en el PTIR, a más tardar el 19 de diciembre de 2014, y estar acompañadas de “…los factores a tener en cuenta para hacer uso de estas en los procesos de toma de decisión” (p. 57).

En el PTIR se presenta claramente la conexión entre las conclusiones y las recomendaciones (estas son presentadas por tipo de recomendación: para el nuevo diseño de atención humanitaria; financieras y operativas), y a cada una se le realiza una descripción y una presentación de su objetivo, se describe su nivel de incidencia esperado, y se puntualizan los aspectos a tener en cuenta para facilitar su implementación. Todo esto satisfizo los componentes del estándar de oportunidad, pertinencia y utilización, que son reglas de alcance.

5.2.2 Reglas de frontera y participantes

En la situación de acción de diseño de la evaluación se contó con la participación de la UARIV, del Grupo de Proyectos Especiales, el cual se encarga del tema de víctimas dentro del DNP, y la DSEPP. Casi la totalidad de los actores que se establecen como encargados de la implementación de recomendaciones participaron desde etapas tempranas del proceso de evaluación, el único actor faltante es el Ministerio Público (Procuraduría General de la Nación, Defensoría del Pueblo y personerías municipales).

Sin embargo, este último tipo de actor, al no tener competencias de planeación dentro de la política en cuestión, su ausencia no representa un mayor problema. Por su parte, el Comité de Seguimiento fue conformado por la UARIV (Oficina Asesora de Planeación, Dirección de Gestión Social y Humanitaria y la Subdirección General), y el DNP (Grupo de Proyectos Especiales y DNP)[57]. Por esto, todos los componentes de los estándares de participación de los actores relevantes, enfoque asociativo y consideración de evaluación conjunta fueron satisfechos en este proceso de evaluación.

[57] Esto con base en la respuesta a derecho de petición del DNP N° 20193600428271.

El contratante de la consultoría fue la DSEPP, sin embargo, los recursos del contrato no fueron puestos por el DNP. Para este contrato se utilizaron recursos del Proyecto Fortalecimiento de la Información Pública Seguimiento y Evaluación para la Gestión por Resultados en Colombia – MONITOREO Y EVALUACIÓN M&E (Contrato de Prestamo: BIRF 7620CO). Aun así, la participación del Banco Mundial se limitó a aprobar y realizar los pagos al consultor una vez eran aprobados por el DNP.

En este caso, la contratación tuvo la modalidad de concurso de méritos con sistema de precalificación con presentación de una propuesta técnica, también llamado "lista corta" (DNP, 2012, p. 2). Para esto, el DNP invitó a firmas consultoras elegibles a expresar su interés en prestar los servicios requeridos hasta el 13 de mayo de 2014 (Banco Mundial, 2014, p. 2). Para el 29 de mayo de 2014 ya se contaba con la lista corta conformada por seis firmas y consorcios.

Para la conformación de la lista corta, los postulantes debieron presentar documentos habilitantes de naturaleza jurídica, financiera, y técnica. En los estudios previos se aclaró el perfil académico y profesional del equipo de trabajo requerido. Dentro de estos requerimientos se destaca: director de proyecto; experto en población desplazada; experto en análisis cualitativo; y experto en análisis cuantitativo (DNP, 2014c, p. 64).

Dependiendo de los años de experiencia y del grado de escolaridad de los miembros del equipo de trabajo propuesto, se asignaba un determinado puntaje a la propuesta (DSEPP-DNP, 2012c, p. 44). Otros criterios de selección del consultor fueron: claridad conceptual, desarrollo metodológico, y herramientas de recolección de información dentro del enfoque metodológico presentado; recurso humano; experiencia de la firma; y apoyo a la industria nacional (p. 34). Con todo esto, se satisfizo el estándar de equipo de evaluación.

También existieron requisitos financieros habilitantes dirigidos a garantizar la capacidad de los proponentes de asumir compromisos y la posibilidad de cumplir con las obligaciones del contrato. Entre estos requisitos estuvieron los indicadores de capital real del proponente, liquidez, nivel de endeudamiento, capital de trabajo, e indicador de crecimiento (pp. 17-18). El valor del contrato fue de $672.579.000, subdivididos en cuatro pagos por $134.519.000 que se realizaron a partir del recibo a satisfacción de los productos. Esto satisfizo el estándar de recursos.

5.2.3 Reglas de elección y acciones

La revisión documental refleja la existencia de una adecuada definición de la intervención al interior de los documentos producidos en la etapa de diseño de la evaluación, particularmente al interior de los estudios previos (pp. 1-3). En el caso de la Política de Atención Humanitaria su definición no conlleva un ejercicio de mucha complejidad, esto es así porque para 2014 ya estaba definida en detalle al interior de la Ley 1448 de 2011 y en el Decreto 4800 de 2011, y porque al observar

esta normatividad, se puede dimensionar que la atención humanitaria es sólo una particular dimensión dentro del componente de “Atención y asistencia”, dentro de la Política de Atención y Reparación Integral a las Víctimas del Conflicto Armado.

Al ser un subsistema que puede describirse como pequeño, la Política de Atención Humanitaria presenta una baja complejidad en su cadena de valor, razón que facilitó que fuese descrita satisfactoriamente, que sus resultados (en las etapas de insumos, procesos y productos a corto plazo de la política) puedan ser descritos como verificables, y que esta descripción tuviera en cuenta los planes y políticas nacionales y locales. Con esto, todos los componentes de los estándares de evaluabilidad y de coordinación y alineamiento estuvieron satisfechos.

En cuanto a la selección de la metodología, en el documento de la convocatoria se aclara que dentro de los servicios de consultoría se incluirá la selección de una metodología de evaluación de operaciones, y el diseño y aplicación de los instrumentos de recolección de la información (Banco Mundial, 2014, p. 1). Adicionalmente en el contrato (DNP, 2014c, pp. 45-54) se especifica a mayor profundidad los lineamientos metodológicos que el consultor debió tener en cuenta, entre estos están:

- Especificar metodología de evaluación de operaciones.
- Definir indicadores.
- Definición del objeto de observación.
- Diseño del muestreo.
- Cálculo del tamaño de la muestra.
- Definición de la estrategia de estimación inferencial.
- Estimación de varianza.
- Descripción de las variables objeto de análisis.
- Marco de análisis (aproximación teórica).
- Instrumentos de recolección de datos cualitativos y cuantitativos.
- Plan detallado de trabajo de campo.
- Realizar una prueba piloto de recolección de datos y realizar los ajustes pertinentes que deriven de esta.

Lo anterior significa que, en contraste con la tendencia general de limitar las decisiones metodológicas del consultor al diseño de la metodología de recolección de datos, en este proceso de evaluación el evaluador tuvo un mayor protagonismo en la definición de la consultoría, a pesar de la revisión y modificación que se hizo a los productos luego de su presentación al Comité de Seguimiento.

Esta metodología busca abordar criterios de evaluación centrados en eficiencia, al direccionar la metodología hacia la determinación de la mejor forma de entregar la atención humanitaria (p. 45); Sostenibilidad, al preguntar por la sostenibilidad fiscal de la política (p. 43); pertinencia, con respecto a todo el componente que se pregunta acerca del nuevo modelo de atención humanitaria; y eficacia, al abordar cuestiones acerca del cumplimiento de las competencias por parte de las entidades

territoriales (p. 41), sobre la suficiencia de los recursos para garantizar la subsistencia mínima, entre otras. Con esto, pueden considerarse satisfechos los estándares de selección del enfoque y método, y de selección de criterios de evaluación.

Dentro de los componentes que se utilizaron a modo de objetivos específicos al interior del contrato de consultoría, se incluyó todo el conjunto de cuestiones a responder (pp. 40-44). En total, se realizaron 45 preguntas orientadoras. A su vez, estas preguntas guardan coherencia con la metodología seleccionada, lo que cumple los componentes del estándar de cuestiones a responder. Con respecto a su respuesta, al observar el informe final se encuentra que este estructuró sus capítulos a partir de los componentes de la evaluación, y al interior de estos, se ubicaron subtítulos que inician con la presentación de la pregunta orientadora que fue incluida en el contrato de evaluación.

A pesar de esto, no se encontró ni en el informe final, ni en el PTIR, respuesta a las siguientes preguntas de evaluación que se incluyeron en el contrato: ¿Es la entrega de atención humanitaria en efectivo coherente con los objetivos de la política de atención humanitaria? ¿Es necesario realizar distinción en el monto de la atención humanitaria con relación al lugar en el que se entrega? ¿Cómo se puede llegar a la población que actualmente no puede acceder a la institucionalidad para la solicitud de atención humanitaria? y, ¿Cuál sería la manera de eliminar los incentivos perversos, en caso de existir, que se presentan a la hora de solicitar la atención humanitaria? Por esto, el estándar de respuesta a las preguntas evaluativas únicamente puede ser considerado como parcialmente satisfecho.

Tanto en el PTIR como en el informe final se realiza un esfuerzo en aras de que el análisis se presente de una manera clara. Los hallazgos son presentados a lo largo de los capítulos del informe final, para las conclusiones se destina el último capítulo del mismo, y al interior del PTIR se presenta una matriz (Econometría y SEI, 2014a, p. 28), en donde se clasifican las recomendaciones a partir del objetivo específico de la evaluación y la conclusión de la que surge, por lo que el estándar de claridad del análisis fue satisfecho.

También se construyeron el resumen ejecutivo (pp. 8-27) y la ficha técnica de la evaluación. El primero contiene una descripción detallada de la política, una síntesis de los principales factores críticos de riesgo y de éxito de los procesos de esta, y se termina con la presentación de conclusiones divididas entre las referidas al diseño, a la financiación y a la operación de la política evaluada. Para la ficha técnica se incluyó descripción de la política, descripción de beneficiarios, objetivo de la evaluación, metodología de la evaluación, principales conclusiones y recomendaciones, por lo que la ficha parece ser mucho más útil en términos de difusión de la evaluación. Con estos dos productos se satisfizo el estándar de claridad y representatividad del resumen.

Entrando en los estándares que deben reflejarse en el informe de la evaluación, se hace una descripción del contexto de desarrollo en el capítulo 6 del informe final “Análisis del modelo actual desde la perspectiva de los hogares” (Econometría y SEI, 2014a, pp. 196-206), y del contexto institucional y político, el cual se presenta en diversas partes del documento, pero particularmente en la introducción y en el resumen ejecutivo. Con lo que el estándar de contexto de la intervención fue satisfecho.

El estándar de lógica de la intervención también presenta satisfacción en todos sus componentes. Dentro del informe final se avanza notablemente en la descripción a profundidad de la intervención, en cada capítulo de presentación de hallazgos se realiza un apartado particular en el que se describe cada uno de los componentes de la política y se presenta en un esquema cada uno de los procesos, los cuales son: proceso de declaración y registro; atención inmediata en dinero; atención inmediata en especie periódico y por evento; atención humanitaria de emergencia y transición. Estos esquemas descriptivos también son presentados en el resumen ejecutivo al inicio de la evaluación.

En cuanto a la metodología, dentro del informe final se establece que se utilizaron tres herramientas (Econometría y SEI, 2014a, pp. 33-34): caracterización del flujo de proceso, con el cual se caracterizó las secuencias de actividades y sus cambios; análisis de flujo de la información, que permitió observar las entradas y salidas de los procesos para identificar deficiencias en los insumos de información; y factores críticos de éxito, que identifica un conjunto de factores que al ser potenciados garantizan el logro de los objetivos de la política. La metodología se presenta de manera sintetizada en un diagrama (p. 35), en el cual se evidencian los insumos que requerirá la consultoría, los procesos que se realizaran para la elaboración de cada uno de los productos requeridos, y los outputs de cada proceso. Esto satisfizo el estándar de explicación de la metodología.

Las fuentes de información son especificadas en primera instancia dentro del contrato, en este se establece que se deben usar: encuestas, grupos focales, entrevistas, registros administrativos y estudios de caso (DNP, 2014c p. 47). Se requiere también una muestra cuantitativa mínima de 1840 hogares (p. 47), número necesario para tener un error máximo del 5% con un universo de 1.035.453 hogares. Y que dentro del diseño del muestreo se utilicen dos estratos: “1) Los hogares que al momento de la consulta hayan recibido atención humanitaria y 2) Los hogares que no han recibido atención humanitaria pero sí la han solicitado” (p. 14).

En la ficha técnica se aclara la aplicación de las herramientas de recolección de datos según el tipo de fuente de información (Econometría y SEI, 2014b, p. 3), a hogares de beneficiarios se les realizó 1854 encuestas y 17 grupos focales; a actores territoriales se dividió en 42 encuestas a alcaldías, 35 encuestas a funcionarios territoriales de la UARIV, 41 encuestas a líderes de mesas de participación, 8 grupos focales con actores territoriales, y 41 entrevistas a los mismos; con actores nacionales se realizaron 2 talleres y 19 entrevistas; también se utilizó información secundaria de bases de datos y documentación relativa a

normatividad, jurisprudencia, documentos de procesos, entre otros. De esta manera, se satisfizo el estándar de validez y fiabilidad de las fuentes de información. También el estándar de consulta y protección de implicados se satisfizo al tener en cuenta la siguiente regla al interior del contrato:

> "El Consultor y su Personal, excepto previo consentimiento por escrito del Contratante, no podrán revelar en ningún momento a cualquier persona o entidad ninguna información confidencial adquirida en el curso de la prestación de los servicios (…) Todos los planos, dibujos, especificaciones, diseños, informes, otros documentos y programas de computación presentados por el Consultor en virtud de este Contrato pasarán a ser de propiedad del Contratante" (DNP, 2014c p. 26)

También se reconocen los límites metodológicos al interior de este informe, entre estos se destacan: la localización de la población relevante, dado su alto grado de vulnerabilidad y movilidad (p.12), la ubicación de hogares que han recibido atención inmediata, puesto que para recibirla no deben estar inscritos en el RUV (p. 12), el alto nivel de pérdidas en los contactos de las unidades de observación con las bases de datos disponibles por deficiencias en la información de contactos (p. 36), con lo que el estándar de mención de límites y cambios de la metodología quedó satisfecho. Al igual que en los anteriores procesos de evaluación, en el informe final no se incluyeron menciones acerca de discrepancias entre miembros del equipo evaluador, por lo que este estándar no se cumple.

De acuerdo con el otrosí añadido al contrato de consultoría (firmado el 11 de diciembre de 2014), el cual modifica el calendario de ejecución del contrato, y establece que su duración irá desde el 19 de agosto de 2014 hasta el 31 de diciembre de 2014. El informe final tiene fecha del 19 de diciembre de 2014 y el PTIR del 26 de diciembre de 2014, por lo que puede considerarse satisfecho el estándar de realización de la evaluación dentro del plazo y el presupuesto previsto.

5.2.4 Reglas de agregación y control

Dentro del contrato de la consultoría se establece que todos los productos serán modificados por parte del consultor, de acuerdo con los comentarios y discusiones que se realicen en el marco del Comité de Seguimiento. Únicamente hasta que el Comité hubiese verificado la inclusión de los ajustes, el consultor podría enviar el producto al supervisor del contrato para proceder con el recibo a satisfacción y con el pago correspondiente (DNP, 2014c, p. 59).

En los estudios previos establecen reglas de agregación adicionales como la responsabilidad del consultor de garantizar el cumplimiento del contrato a partir de la constitución de una póliza de garantía de seriedad de la oferta (DSEPP-DNP, 2012c, p. 33); el plazo de una semana que tiene el consultor para presentar los productos una vez haya recibido los comentarios del Comité de Seguimiento (p. 58);

y que la Secretaría Técnica del Comité de Seguimiento (DSEPP) se encargó de recibir los comentarios de los otros miembros del Comité, consolidarlos y depurarlos, para entregarlos al consultor (p. 56). Con lo anterior, los estándares de control de calidad y de inclusión de observaciones de los implicados fueron satisfechos completamente.

5.2.5 Reglas de posición

Frente a la tensión que existe entre las reglas de agregación referentes al control ejercido por parte de la DSEPP, con las reglas de posición que se centran en la independencia del evaluador y de un proceso evaluativo libre y abierto, se debe decir que no se encontraron reglas de posición que estén direccionadas a fortalecer o a direccionar en algún sentido estos dos estándares, por lo que el establecimiento de las reglas de control mencionadas en el anterior subtítulo, restringen fuertemente la posibilidad de un juicio independiente por parte del evaluador.

El equipo evaluador tuvo un escenario de discrecionalidad correspondiente al desacuerdo que puede manifestar con respecto a los comentarios del Comité de Seguimiento. Sin embargo, el proceso de elaboración de los productos que son descrito en sus conceptos técnicos refleja que estos comentarios fueron incluidos por parte del equipo evaluador sin mayor resistencia. En cuanto a su naturaleza, los siguientes ejemplos ayudan a caracterizarla:

> “Desarrollar el capítulo 6 del documento ya que la mayoría de los apartados presentan cifras sin describir y analizar, lo cual lleva a que en un número importante de casos no se respondan las preguntas orientadoras” (DSEPP-DNP, 2014e, p. 3)
>
> “Analizar la pertinencia de las afirmaciones teniendo en cuenta los avances en gestiones sobre algunos procesos, así como los avances en aspectos que no se tuvieron en cuenta debido a que, en el momento de estar en campo, no estaban aún en implementación” (p. 3).
>
> “Revisar las afirmaciones presentadas, debido a que algunas de ellas son realizadas sin tener en cuenta toda la información suministrada por la UARIV, como es el caso de la entrega o suspensión de la AH, la cual no está basada exclusivamente en la temporalidad, sino que contempla la evaluación de las carencias de las víctimas” (p. 3).
>
> “El nuevo modelo no se contempla en las recomendaciones, muchas van dirigidas a cuellos de botella del modelo actual. No se contrasta modelo anterior con el nuevo de forma completa e integral. Se trabajan algunos puntos, pero de manera aislada en recomendaciones específicas” (DSEPP-DNP, 2014f, p. 3).

> “Se muestran consideraciones para la implementación del nuevo modelo, pero no hay un plan integral de recomendaciones de las operaciones del nuevo modelo” (p. 3).
>
> “Falta desarrollo técnico de las recomendaciones” (p. 3).

En general los comentarios se direccionan a la agregación de temas pasados por alto por parte del equipo evaluador, y al perfeccionamiento de algunos elementos. Sin embargo, es clara la intencionalidad que muestra el Comité de Seguimiento de focalizar las recomendaciones de la evaluación hacia la implementación del nuevo modelo, en contraposición a su posible modificación con base en la evaluación. Con todo esto, solo pude considerarse como parcialmente satisfechos los estándares de independencia y de proceso evaluativo libre y abierto.

5.2.6 Reglas de información

En cuanto a la documentación acerca de la finalidad y la expectativa, dentro del SECOP se pudo acceder a los siguientes documentos: Estudios previos; Solicitud de expresiones de interés para la conformación de la lista corta; Anexo técnico que contiene los requerimientos y actividades específicas que debe realizar el consultor. A pesar de que no se encontró el pliego de condiciones, en estos documentos existe la suficiente información para poder considerar que se cumplió el estándar sobre documentación acerca de la finalidad y la expectativa.

También se observa que el informe de la evaluación tiene unos requerimientos claros al interior de los estudios previos y el contrato. En general, se puede decir que el informe de la evaluación tiene una buena estructura que lo hace fácilmente comprensible, sin embargo, dentro de los lineamientos del contrato se encuentra la inclusión de la respuesta a cada una de las preguntas orientadoras establecidas en todos los componentes de la consultoría (Econometría y SEI, 2014a, p. 56). Como se observó, las respuestas del componente correspondiente a las recomendaciones, no se incluyeron en el informe final, sino en el PTIR, por lo que no se cumplió totalmente los requerimientos al informe final que se encuentran en el contrato relativos al informe final. Esto significa una satisfacción parcial del estándar de informe de la evaluación.

Adicionalmente, se debe considerar que la no inclusión de estas respuestas a las preguntas relativas a las recomendaciones dentro del informe de evaluación limita el alcance de la difusión externa que se ha hecho de los productos de la evaluación, esto es así porque el PTIR (que es donde se incluyeron las respuestas a este componente) no es un producto que se difunda a través del portal de SINERGIA. En cuanto a difusión externa, para este proceso de evaluación se difundió la ficha técnica de la evaluación y el informe final (el resumen ejecutivo está dentro del informe final de la evaluación).

Para difusión interna, además de los productos ya mencionados, dentro del contrato se requieren dos presentaciones: una extendida y otra simple (p. 59). La primera incluyó una presentación extendida de la metodología, el alcance de la evaluación, así como de todos los resultados y recomendaciones. La simple incluyó una descripción corta de la metodología, el alcance, y los resultados y recomendaciones más relevantes, cuya priorización se determinó en discusiones con el Comité de Seguimiento a la evaluación. Según el PTIR, la difusión de estos productos se enfocó a los siguientes actores (Econometría y SEI, 2014c, p. 29): actores institucionales involucrados en la aplicación de los resultados de la evaluación al interior de la UARIV y otras entidades con las que esta se coordina y con representantes de las mesas de participación de víctimas. Quizás la norma más importante con respecto a la difusión consistió en difundir los productos a los directivos de las entidades que conformaron el Comité de Seguimiento (DNP, 2014c, p. 58). A pesar de la limitación en la difusión externa, puede considerarse satisfecho el estándar de difusión de la evaluación.

5.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la Política de Atención Humanitaria

Al verificar el análisis de actores que se realiza en el PTIR (Econometría y SEI, 2014c, p. 28), y se contrasta con la información de las entrevistas, se puede establecer que se hallaron y se involucraron a todos los usuarios previstos en todo el proceso de la evaluación, al incluirlos dentro del Comité de Diseño y en el Comité de Seguimiento a la evaluación. Por lo general, los actores a los cuales se les identificó su responsabilidad de implementación de las recomendaciones de la evaluación son unidades organizacionales al interior de la UARIV. Puntualmente, al observar la lista de funcionarios miembros del Comité de Seguimiento que se indica al inicio del informe final (p. 2), se tiene que participaron funcionarios de la Dirección de Gestión Social y Humanitaria (UARIV); Oficina Asesora de Planeación (UARIV); Grupo de Proyectos Especiales (DNP); y DSEPP. Frente a las reuniones que tuvo el Comité de Seguimiento con el equipo de evaluación se pudo caracterizar:

> "[la coordinación] fue fluida. Creo que tiene ver con que Econometría es una consultora que ya había trabajado con la Unidad para las Víctimas, conocíamos a algunas de las personas que trabajaban allí, respetamos su criterio técnico. Ellos se ganaron la convocatoria por concurso abierto entonces había ya un flujo de información y comunicación entre nosotros y la consultora y había una disposición muy alta desde la Unidad en ese momento de hacer la evaluación, porque finalmente fue la Unidad quien decidió priorizar esta evaluación, habían todas las voluntades para que la evaluación ocurriera, no era una evaluación impuesta donde los resultados pudieran jugar en contra de nosotros, o fueran percibidos como una amenaza, al contrario era algo que nosotros queríamos sacar rápido para poder validar el nuevo decreto. Yo recuerdo bien el momento, fue fluida, fue una evaluación

> que yo creo que funciono de manera armoniosa" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Como se observa, la percepción que quedó del equipo de trabajo conjunto es positiva. También se puede decir que el DNP tuvo en cuenta las demandas de información de los usuarios previstos al momento de determinar el alcance de la evaluación y las preguntas orientadoras, esto se refleja particularmente en la inclusión del componente de la evaluación dirigido hacia la evaluación de la pertinencia del nuevo modelo: "… si hubo una receptividad grande del DNP, en definir un alcance no solamente que no solamente evaluara hacia el pasado, sino que también tuviera en cuenta una validación de lo que estamos pensando hacia futuro a partir de lo que encontraran en la evaluación" (Asesor de formulación de la Política de Atención Humanitaria, UARIV). Esto refleja un trabajo fluido del equipo de trabajo conjunto con respecto a los usos y a las preguntas de evaluación.

Al preguntar acerca del grado en que el DNP y el equipo evaluador tomaban en cuenta las preocupaciones o inquietudes al momento de tomar decisiones alrededor de la evaluación, es decir, su grado de reactividad, se contestó:

> "Si, totalmente (...) en ese Comité, por lo general, la revisión de la UARIV de los productos de las entregas del consultor, son tenidas en cuenta casi como un aval para seguir adelante con el proceso de la consultoría, entonces en ese sentido, y no hablo solamente de la evaluación de atención humanitaria sino otras que había en el marco de evaluación de políticas públicas, el DNP siempre fue muy respetuoso con que la UARIV realmente sea quien valide las entregas parciales del consultor. Entonces yo creo que en ese sentido, si recuerdo que había reuniones periódicas para evaluar, para revisar los productos que entregaba el consultor, y siempre tenía que haber un aval, o como un visto bueno de revisado de lo que entregaran por parte de la UARIV, y que si eso implicaba hacer ajustes, darle alcance, completar, cambiar, siempre se tenían en cuenta, entonces yo recuerdo mucho que esa relación siempre fue en doble vía, tanto entre la UARIV y el consultor, como entre la UARIV y el DNP, y el DNP con el consultor, entonces, lo recuerdo de manera positiva, o sea si se tenían en cuenta (…) sé que también pasó de una manera similar en la evaluación que se hizo de Reparación Colectiva y de Retornos y Reubicaciones, sin yo estar directamente como miembro de esos comités, pero sé que era la dinámica de trabajo con el DNP y con los consultores, y era tener siempre en cuenta la retroalimentación de la UARIV en la revisión de los productos parciales que se entregaban como parte de la evaluación" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Se debe resaltar la percepción del elevado grado de reactividad que tiene el DNP frente a las necesidades de los usuarios previstos, en los distintos procesos de evaluación en los que participa la UARIV, esto ayuda a caracterizar esos otros dos

procesos, lo cuales también se tuvieron en cuenta para el presente estudio. Esta reactividad del DNP hacia las necesidades de la UARIV se observa extendida en la relación entre el DNP y el equipo evaluador. En el contrato se estableció: “Los Consultores deben otorgar máxima importancia a los intereses del contratante, sin consideración alguna respecto de cualquier labor fututa” (DNP, 2014c, Pág. 24).

Con respecto a los debates metodológicos que deben darse en el equipo de trabajo conjunto, se encontró información acerca de la realización de talleres de los tres actores en donde se identificaron procesos y factores críticos a abordar, con lo que posteriormente, se construyeron los indicadores e instrumentos de recolección de información (Econometría y SEI, 2014a, p. 34). Adicionalmente, todas las acciones correspondientes a la metodología de la evaluación, en general, debieron tener aprobación previa por parte del ANP antes de ser realizadas por el evaluador-consultor (p. 26; p. 54).

Una vez debatida y acordada, la metodología no necesitó grandes cambios a lo largo de la ejecución del contrato, esto es probable que se deba a la corta duración de este, lo que reduce las probabilidades de cambios en las necesidades de los usuarios previstos. Al inquirir al respecto, se contestó:

> “No recuerdo si hubo necesidad de ajustar esa metodología una vez se acordó, pero creo que la forma como se acordó inicialmente tuvo resultados como lo esperábamos (...) No creo que hubo necesidad de ajustar la metodología, no sé si en el camino hubo algún ajuste menor que hayamos hecho, producto de pronto de no haber logrado recoger la información como se esperaba, yo creo que no pasó, creo que más bien lo que sucedió es que la metodología la pensamos bien desde un comienzo” (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

En las últimas etapas de la evaluación, el equipo de trabajo conjunto sirvió para el involucramiento de los usuarios previstos, tanto en la interpretación de los hallazgos, como en la construcción de las recomendaciones. Además de la inclusión de todos los comentarios realizados por parte del Comité de Seguimiento, algunos de los cuales se enfocan en la modificación de las valoraciones y las recomendaciones del consultor (tal como se vio en la presentación de las reglas de posición), se observa que involucró a los usuarios para la determinación acerca de lo que es significativo, y para deliberar acerca del costo, los beneficios y los desafíos de implementar determinadas recomendaciones:

> “Yo creo que si hubo una incidencia. Por un lado es cierto que fueron unos validadores de que lo que estábamos formulando desde la intuición que era algo que podía funcionar, sin embargo, se abrieron discusiones interesantes por ejemplo las de prestamo a la Banca Multilateral, visitas in situ, recomendaciones que vienen de externo y que podrían parecer lógicas ante cualquier ciudadano desprovisto de un sesgo. Nos llevaron a discusiones interesantes de si estábamos cometiendo un error al no hacer visitas in situ, si hacer encuestas telefónicas y encuestas desde la oficinas tenían un sesgo,

> de ahí en adelante si nos cuestionamos mucho, pensamos mucho en hacer evaluaciones que a partir de la vivienda nos mostraran si existían o no diferencias estructurales en tres la información que usted levanta en una visita domiciliaria y la información que usted levanta en una encuesta telefónica" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Por último, frente a la coordinación que debe realizarse para el seguimiento a la implementación de los hallazgos de la evaluación, las entrevistas reflejan que se realizaron algunos ejercicios de consulta por parte del Grupo de Proyectos Especiales del DNP, quien tiene la función de coordinar los ejercicios de planeación en los temas de víctimas y posconflicto al interior de esta entidad (es decir, juega el papel de una dirección sectorial del DNP). Entonces, se tiene que, en el marco de ejercicios como la elaboración del nuevo Plan Nacional de Desarrollo, o la aprobación de los presupuestos anuales de inversión, este grupo realizó preguntas a la UARIV acerca de si se tuvieron en cuenta las recomendaciones de las evaluaciones:

> "… digamos que en los ejercicios que hace el DNP, por ejemplo, para Plan Nacional de Desarrollo, para efectos de aprobación de los presupuestos cada año, generalmente de la recepción que ellos tienen, y las preguntas que hacen es si se tuvieron en cuenta las recomendaciones de las evaluaciones que se hacen (…) nos permitió a todos dentro de gobierno, como sustentar la necesidad de los cambios en la forma como se venía implementando esa política, entonces claramente cuando hubo, por ejemplo, ejercicio presupuestal 2016, 2017, 2018, siempre si se preguntaba desde el DNP, venga ustedes en la evaluación ¿qué están haciendo en cálculo del presupuesto? ¿están teniendo en cuenta los ajustes que surgieron de la evaluación que se hizo del proceso?" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Sin embargo, debido a su insipiencia y simplicidad, no puede considerarse un ejercicio de seguimiento sistemático y, por ende, su capacidad de incidencia en el uso de la evaluación es limitada. Lo anterior, sumado a la terminación del contrato de consultoría, permiten establecer que tampoco existió una búsqueda proactiva del uso.

5.4 Análisis de la efectividad de la evaluación de la Política de Atención Humanitaria

La evaluación produjo ocho recomendaciones. La primera corresponde a actualizar la información de registro, a partir de coordinar con la estrategia de Gobierno en Línea, la ampliación del Formato Único de Declaración Digital en todos los

municipios, luego de esta ampliación, se debería establecer como obligatorio su uso (Econometría y SEI, 2014c, p. 24). Con esto, se podría contar con información actualizada de contacto e identificación, y garantizaría que los hogares se enteren y cobren su giro de atención humanitaria (p. 30). Al preguntar acerca del estado de implementación de esta recomendación se manifestó:

> "Esta es una recomendación que igual siempre ha sido parte de la estrategia de registro, de mejorar la toma de declaración. Dentro de la Unidad para las Víctimas el área encargada de tomar las declaraciones es una dirección que se llama Dirección de Registro y Gestión de la Información, y ellos se inventaron el Formato Único de Declaración Digital para reemplazar el formato escrito (…) La recomendación no es una recomendación que se haya inventado la evaluación, es decir la toma digital de la declaración ya existía en el momento en el que se hizo la evaluación pero estaba en un proceso incipiente, estábamos hasta ahora arrancando a mejorar la conectividad de las personerías, de las alcaldías, haciendo campañas para que se aumentara el uso del Formato Digital, entonces es una recomendación que se hizo en un momento en el que ya el proceso estaba arrancando, ya la Unidad para las Víctimas estaba tomando declaraciones digitales y ese proceso de capacitación, de mejorar la conectividad de los puntos donde se toman las declaraciones en las alcaldías ha sido un proceso continuo de mejora en la Unidad, yo no tengo la estadística de en este momento del 100% de las declaraciones cuántas se están tomando digital, para compararlo con el porcentaje que se tomaba en el 2014 cuando se hizo la evaluación, pero efectivamente esta recomendación que venía ya aplicándose y se ha seguido aplicando" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Al haberse tomado la decisión previamente a la obtención de los hallazgos de la evaluación, y a la construcción de la recomendación, la implementación de esta última no significó un aporte o mejora del proceso de toma de decisiones, únicamente ayudó a "validar" (argumentar, justificar y/o legitimar) decisiones ya tomadas. En este sentido, la implementación de esta recomendación no reflejó efectividad del proceso de evaluación.

La segunda recomendación sugirió verificar "in-situ" la caracterización de la población beneficiaria de la atención humanitaria (Econometría y SEI, 2014c, p. 25). Concretamente, se propuso incorporar al Plan de Atención, Asistencia y Reparación Integral (PAARI), un procedimiento de actualización y verificación en campo, por medio de visitas aleatorias y focalizadas a dos tipos de hogares: los que muestran una superación de sus carencias; y a aquellos en los cuales la información administrativa resulte muy diferente de la información que reportan por medio de la encuesta telefónica o presencial. Al respecto se comenta su escaso grado de factibilidad:

> "… a hoy la Unidad para las Víctimas no hace verificaciones in situ (…) al momento en el que se estaba haciendo esta evaluación una recomendación

> fácil sería decirle a la Unidad para las Víctimas 'si usted quiera verificar si realmente el que le está pidiendo la plática la necesita, pues usted debe ir a verificar si esta persona vive bien, si come bien, para determinar si necesita o no la ayuda', pero cuando usted dice como entidad de gobierno, que tiene un presupuesto limitado, que además tiene una capacidad operativa limitada y le dicen 'vaya y visite a 1.200.000 hogares para verificar que comen bien, los materiales del piso, los materiales de las paredes, el hacinamiento, el acceso a servicios públicos, haga una visita de esas que son realmente exhaustivas' (…) eso es una operación que no existe en Colombia (…) En este caso lo que recomienda la evaluación es apuntarle a una visita in situ tipo censal, que en la práctica es inviable, entonces lo que nos decían era 'intenten por lo menos seleccionar algunos hogares donde ustedes puedan ir a verificar de manera directa en el hogar si efectivamente necesitan o no la ayuda' (…) en Colombia el fenómeno de desplazamiento hizo que las personas huyeran de sus lugares de residencia habitual y se ubicaran por todo Colombia , entonces encontramos desplazados en toda Colombia, una alta concentración en zonas urbanas y dentro de las zonas urbanas no son fácilmente identificables (…) Entonces la estrategia que pensamos que era más óptima en términos de recursos y de tiempo, era aumentar el intercambio de información ¿qué significa esto? existen otras entidades del gobierno que hacen visitas domiciliarias, o que levantan información de caracterización población desplazada, entonces si ya hay entidades del gobierno que ya están haciendo la tarea de recoger información ya sea a través de visitas, para aquellas personas que son beneficiarias de ciertos programas, o a través de encuestas que ellos diligencian cuando las personas se inscriben a sus programas, por qué no echarle mano a esa información" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Con esta recomendación se buscaba corroborar información suministrada por medio de encuesta, y mejorar la fortaleza de los argumentos en caso de eventuales acciones judiciales contra decisiones tomadas con esa información (Econometría y SEI, 2014c, p. 30), como el retiro de la ayuda a las personas que se considere apropiado. Ante esta necesidad, se optó por otra solución distinta a la que sugirió la recomendación de la evaluación, por lo que esta última al no ser implementada, no refleja efectividad de la evaluación.

La tercera recomendación consistió en el establecimiento de un procedimiento progresivo de salida de los beneficiarios, en donde el monto percápita y el plazo de la atención se calcule en función de la probabilidad de superación sostenible de carencias de subsistencia mínima. Para esto era necesario realizar un cálculo de una función denominada "puntaje de avance en la transición" la cual manifestaría la probabilidad entre 1 y 100. Con esto, los montos y el plazo entre giros podrían calcularse dependiendo del puntaje y del avance de este (p. 25). Con esto se buscó un tránsito suave de los hogares hacia otras medidas de atención, asistencia y reparación (p. 30). Al respecto se aclaró:

> “Esa recomendación no se tuvo en cuenta (...) les dijimos que en el momento en que nosotros identifiquemos que alguien ya vive en condiciones óptimas para efectos de la entrega de atención humanitaria, que tiene una frecuencia y un acceso para el consumo de alimentos óptimos, de acuerdo a unos estándares, le suspendemos la ayuda, es decir le cortamos la entrega del beneficio. Entonces el DNP dijo ‘que tal que esa persona este accediendo a alimentación, tenga cierta calidad en su vivienda, gracias a la ayuda que usted le está entregando, entonces si usted le quita esa ayuda, pues la persona va a caer nuevamente en una carencia habitacionales y de ocupación’; ellos nos decían que ‘mejor un desmonte gradual’, entonces en lugar de decirle usted ya vive bien y come bien, entonces yo le corto el beneficio y chao, ellos decían que ‘mejor no hacer eso tan tajante, mejor de manera progresiva’, tenía lógica, es una forma de verlo, pero en la práctica nosotros no lo hicimos (…) lo que hicimos fue establecer un procedimiento administrativo para que la persona a la que se le suspendiera la atención humanitaria se le informara sobre esta decisión, y que como ya accede a ciertas condiciones habitacionales y de alimentación (…) se le informa que puede ir y tocar la puerta de otro programa (…) Lo que nosotros queríamos con todos estos cambios que le propusimos al nuevo modelo era cortar con esa dependencia y con ese asistencialismo en el que veníamos trayendo a las víctimas del desplazamiento (…) hay toda una serie de medidas de carácter social, económico que no se llaman atención humanitaria, llámelas proyectos productivos, proyectos de generación de ingresos, programas de capacitación, mejoramiento de vivienda, que son los programas que tiene el Ministerio de Vivienda, que tenía el INCODER en su momento y ahora las nuevas agencias de tierras, que le permiten a la gente no recibir una plata, sino aplicar a unos proyectos en los que pueden mejorar la calidad de vivienda” (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Esta recomendación tampoco se implementó a pesar de que fue factible, y se consideró válida por parte de los miembros del Comité de Seguimiento, por lo que no refleja efectividad de la evaluación. En cuarto lugar, se recomendó cubrir y prevenir los riesgos judiciales y fiscales a través de un indicador de alerta temprana que identifique situaciones de crecimiento acelerado de demanda de atención humanitaria, y de aceleración de acciones de título, así se priorizaría recursos y se dirigirían a los municipios en donde se identifique este riesgo (Econometría y SEI, 2014c, p. 26). Con esto se buscó solucionar el riesgo de iliquidez por posibles tutelas, lo que permitiría seguir atendiendo oportunamente a la población que no interpone este recurso (p. 30). Al preguntar por su estado de implementación, se respondió:

> “Esto es complejo. Esto es una recomendación que hicieron desde Econometría (…) decían ‘ustedes tienen un problema, y es que no tienen plata suficiente para atender la demanda’ porque es demasiada demanda y en ese momento la valoración que hacíamos a través de registros

> administrativos para determinar si necesitaban o no la ayuda, era muy incipiente, no teníamos información suficiente, no teníamos suficientes registros administrativos, esta encuesta que diligenciamos no tenía un cubrimiento amplio, no teníamos la plata para hacer visitas domiciliarias; entonces lo que estábamos haciendo era entregando una atención humanitaria a ciegas (…) la cuestión es que no se está haciendo una evaluación real de la situación de los hogares, entonces más bien póngase pilas en mirar cómo está la gente y empiece a disminuir la gente que recibe atención humanitaria" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Como se observa, el indicador para cubrir riesgos fiscales y judiciales no se implementó y no refleja efectividad de la evaluación, lo que se hizo fue disminuir el número de personas que reciben la atención humanitaria para reducir la presión fiscal. Esto se hizo a partir del establecimiento de nuevos estándares de subsistencia mínima (modelo de carencias) y del fortalecimiento de la información a partir de la articulación con bases de datos de varias entidades del Estado, tal como se puede observar en las citas anteriores, así como en la que responde a la siguiente recomendación.

La quinta recomendación también se dirigió a reducir la presión fiscal. En este caso se recomendó un financiamiento de corto plazo para garantizar las fechas de pago. A partir del establecimiento de instrumentos financieros que permitan contar con recursos oportunamente, el operador bancario podría realizar el pago en las fechas oportunas para posteriormente cruzar cuentas con la UARIV (Econometría y SEI, 2014c, p. 26). Con esto, se esperaba contar con los recursos para cumplir con las fechas de pago, sin esperar a disponer de recursos por vía de reintegros o adiciones presupuestales del Ministerio de Hacienda. Al respecto se comentó:

> "Esto tampoco es una recomendación que fue factible, lo que decían estos manes era 'digamos que usted no tiene la plata suficiente para pagar las atenciones humanitarias que le piden hoy, pero entonces comprométase con una fecha' (…) lo que hacíamos era darle una priorización a aquellos hogares que tenían una composición más vulnerable (…) el orden de priorización que le dábamos a las personas y ante los cuales si teníamos una fecha aproximada de pago, terminaban en la cola en la medida en la que entraban tutelas (…) eran hogares donde no había realmente una vulnerabilidad asociada al desplazamiento, de acuerdo con la evaluación no deberían están recibiendo la ayuda, pero tenían ya aprendido el mecanismo para colarse a la fila, entonces ante eso no había nada que hacer. El uso del medio de la tutela nos llevó a que no pudiéramos tener una respuesta de fecha exacta de pago para aquellos hogares que si hacían la cola de manera normal (…) Esta recomendación no se pudo aplicar, lo que soluciono este problema fue la puesta en marcha de un modelo de caracterización y evaluación de las necesidades del hogar que llevo a que a partir del 2015, ya empezáramos

> nosotros a informarle a los hogares que no necesitaban la ayuda, que no la iban a recibir. Nosotros arrancamos un modelo nuevo, que es el modelo que llamamos modelo de carencias, modelo de subsistencia mínima que se decretó a través de un decreto presidencial que se firmó en diciembre de 2014, a los quince días que salió la evaluación, porque a medida que nosotros hicimos la evaluación con DNP veníamos diseñando el nuevo modelo, ese modelo termino de diseñarse en agosto de 2014, surgió el proceso de revisión jurídica y en diciembre se firmó un decreto, ese decreto modifica la forma como se entrega la atención humanitaria, y parte de los hallazgos y observaciones que hizo el DNP se tuvieron en cuenta para validar que lo que nosotros estábamos haciendo era la forma correcta (…) a partir del 2015 empezamos a decirle a mucha gente, casi el 40% de las solicitudes arrojaban que las personas ya no necesitaban la ayuda" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Esta cita muestra que no se implementó financiamiento a corto plazo para la reducción de la presión fiscal, y muestra que las decisiones con respecto al mejoramiento de la política pública en términos de financiamiento, mejoramiento de la información y evaluación de las carencias de los hogares, se tomaron hasta agosto de 2014, momento en que se finaliza el diseño del nuevo modelo de atención humanitaria, sin embargo, las recomendaciones del equipo evaluador y corregidas por parte de la misma UARIV, no surgen sino hasta el mes de diciembre, por lo que estas no son tomadas en cuenta, y hace que la evaluación sirva únicamente para legitimar y justificar el nuevo diseño, pero no para mejorarlo (lo cual era el objetivo del componente de recomendaciones de esta evaluación).

En sexto lugar, se recomendó la integración de los componentes de la Atención Humanitaria que se manejan por separado entre la UARIV y el ICBF, a través de una figura jurídica (fiducia o convenio), para brindar una sola atención al hogar y establecer un solo equipo de atención de acciones constitucionales (Econometría y SEI, 2014c, p. 27). Esta recomendación estuvo dirigida a simplificar los trámites, aumentar la eficiencia en la utilización de los recursos públicos, y la liberación de recursos del ICBF para sus actividades misionales (p. 30). Al preguntar acerca de su estado de implementación se respondió:

> "… no se hizo a través de fiducia, lo que se hizo fue modificar un decreto (…) eso lo solucionamos de manera más estructural al modificar la forma como la ley daba las órdenes, entonces a través del decreto que se firmó en diciembre de 2014, se modificó el decreto reglamentario de la Ley de Víctimas y en ese decreto quedo que a partir del 2015 la Unidad para las Víctimas iba a pagar tanto el componente de alojamiento como el de alimentación" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Como se observa, la integración de los pagos se realizó, pero no a través de los mecanismos propuestos por el equipo evaluador. Sucede muy similar a la recomendación anterior, la decisión se tomó con el establecimiento del Decreto

2569 de 2014, este decreto tiene fecha del 12 de diciembre de dicho año, mientras que el PTIR tiene fecha del 26 de diciembre. A parte de la poca oportunidad que tiene la evaluación, se puede deducir que la decisión de integrar los pagos, así como el modo de hacerlo, ya se había tomado al momento de contar con los resultados de la evaluación. Esta recomendación no fue implementada del modo en que fue formulado por el equipo evaluador.

La séptima recomendación consistió en el establecimiento de un programa de pedagogía del pago, en el cual se le brindaría conocimientos acerca del uso de herramientas financieras, de información, de gasto y ahorro a los hogares (Econometría y SEI, 2014c, p. 27). Con esto, se esperaba hacer saber a las personas en dónde y cómo realizar los cobros, y las medidas de seguridad y precaución para su mejor uso (p. 30). Al respecto, se comentó:

> "Digamos que eso no es recomendación valida, se tuvo en cuenta de forma parcial e incipiente, nosotros tuvimos un proyecto con una organización no gubernamental en 2015, que buscaba hacer algún tipo de educación financiera para que las personas que recibían la atención humanitaria pudieran dar una mejor inversión a esos recursos, que realmente entendieran que era [para] pago de arriendo y de mercado, no para pagar deudas o para comprarse un televisor nuevo. Pero fue muy incipiente (…) en términos de protocolo se hizo, se incluyeron los protocolos como información de como se debería gastar esa plata, pero de ahí a que digamos que esa información se haya transmitido a las personas que solicitamos la ayuda en todo el territorio, no estoy seguro. Una campaña grande de pedagogía no se hizo" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Esta recomendación no fue tenida en cuenta. Por último, se recomendó establecer un enlace con el resto de medidas de atención, asistencia y reparación, si bien el nuevo modelo ya consideraba la remisión a otras de estas medidas como parte de las notificaciones del pago de la atención humanitaria, se recomendó "… seguir mejorando la articulación inter e intrainstitucional como la necesidad de informar adecuadamente al hogar, sobre dónde se encuentra en la ruta, sus posibilidades de acceso a otras medidas como la indemnización administrativa y programas de generación de ingreso" (Econometría y SEI, 2014c, p. 27). Con esto se esperaba complementar el tránsito de la atención humanitaria hacia las demás medidas (p. 30). Con respecto a esto, se respondió:

> "La evaluación lo que hizo fue documentarlo, nosotros veníamos ya buen tiempo trabajando en eso, porque para que un decreto salga hay todo un proceso técnico de formulación de la política, que en nuestro caso arranco desde el 2013, entonces nosotros veníamos diseñando todo el proceso, no solamente el de atención humanitaria sino el más amplio que es el de superación de vulnerabilidad (…) Esto es una recomendación que ellos nos hacen, pero nosotros ya veníamos carneándonos que iba a ser parte integral

> de nuevo modelo, ellos documentan, sí, pero no es gracias a ellos que nosotros nos damos cuenta que esto hay que hacerlo, digamos lo que permite la evaluación es reafirmar cosas que usted ya viene pensado. Son muy pocas las evaluaciones que a usted le abren un camino completamente nuevo, lo que hacen estas evaluaciones es validar y actúan como un validador externo, técnico de que lo que usted está pensando como entidad va a funcionar, para nosotros la mayor ventaja de esta evaluación es que en el momento en el que lanzamos el decreto, teníamos un validador técnico nosotros dijimos: 'vea este es el decreto, está sustentando en una evaluación hecha por el DNP en la que dice que efectivamente lo que estamos pensando, es la forma en la que se debe hacer" (Asesor de formulación de la Política de Atención Humanitaria, UARIV).

Entonces, esta última recomendación no apunta a mejorar el nuevo diseño de política que se quería aprobar, únicamente busca ratificar que se debe seguir por el mismo camino que ya se venía pensando. No implicó un mejoramiento del proceso de toma de decisiones dentro de la intervención, puesto que justifica decisiones previas. Se debe subrayar la apreciación de este funcionario, en cuanto a que son "muy pocas" las evaluaciones del DNP que abren "un camino nuevo". Esta percepción está en completa contraposición con las reglas de alcance de la política de evaluación de políticas públicas, particularmente en cuanto a mejorar la toma de decisión.

Como se explicó en el capítulo 1, un mejoramiento de la toma de decisiones como modo de uso de la evaluación implica un cambio en el modelo mental de los usuarios. Este cambio de modelo mental se ve limitado al momento de establecer como finalidad de la evaluación la validación del nuevo modelo que tienen actualmente los usuarios previstos, tal como se observó en el análisis de las reglas de alcance.

5.5 Conclusiones

Con respecto a la satisfacción de los estándares profesionales de calidad, únicamente se tiene problemas parciales en los estándares de respuesta a preguntas evaluativas; independencia del evaluador; proceso evaluativo libre y abierto; e informe de la evaluación. No se cumple ningún componente del estándar de mención de discrepancias entre miembros del equipo evaluador. Aun así, es una de las evaluaciones que mejor satisface todo el conjunto de estándares de calidad.

En cuanto a la coordinación, se tiene que en efecto se hallaron y se involucraron a todos los usuarios previstos en todo el proceso de evaluación. Los usuarios tienen una percepción positiva acerca de la fluidez de la coordinación y la comunicación dentro del equipo de trabajo conjunto, se manifestó que el DNP tuvo en cuenta las demandas de información de los usuarios previstos al momento de determinar el

alcance de la evaluación y las preguntas orientadoras, y que los comentarios de la UARIV eran “casi un aval” para continuar con el proceso, lo que refleja que se tomaron en cuenta las preocupaciones e inquietudes de los usuarios previstos, por lo que se evidenció un elevado grado de reactividad del DNP frente a las necesidades de los usuarios previstos, en los distintos procesos de evaluación en los que participa la UARIV.

La metodología se discutió y se aprobó de manera conjunta, posteriormente no necesitó modificaciones debido a posibles cambios en las necesidades de los usuarios previstos. Así mismo, el equipo de trabajo conjunto sirvió para el involucramiento de los usuarios previstos en la interpretación de los hallazgos y en la construcción de recomendaciones. Sin embargo, el seguimiento a la implementación de los hallazgos de la evaluación fue muy simple e incipiente.

En notorio y muy diciente contraste con este balance positivo, se tiene la calificación más baja posible en cuanto a uso de la evaluación. De ocho recomendaciones formuladas, que buscaban mejorar el nuevo modelo de política propuesto, ninguna refleja la implementación de acciones que surgieran de la ampliación del rango de acciones y resultados posibles, dos de estas no significaron un cambio en el modelo mental sobre la intervención de los usuarios, por lo que sirven de justificación de decisiones previas, y no reflejan un mejoramiento de la toma de decisiones.

Las decisiones con respecto a las acciones a implementar para mejorar la Política de Atención Humanitaria se tomaron en el proceso de formulación del Decreto 2569 de 2014, el cual entró en vigencia el 12 de diciembre de ese año y cuyo proceso de diseño terminó en agosto (según la misma UARIV). Por su parte, los resultados de la evaluación estuvieron disponibles hasta el 26 de diciembre (DSEPP-DNP, 2014f, p. 3). Luego de que se expidiera este decreto, a 2019 no se han realizado modificaciones adicionales a la institucionalidad de la Política de Atención Humanitaria. Esto quiere decir que no hubo apertura por parte de los funcionarios a cargo de la política, de cambiar la institucionalidad de la política una vez se tuvieron disponibles los resultados de la evaluación.

6. Efectividad de la Evaluación de Operaciones y de Resultados de la Política de Retornos y Reubicaciones a las Víctimas de Desplazamiento Forzado

6.1 Presentación de la Política de Retornos y Reubicaciones

La Política de Retornos[58] y Reubicaciones[59] es una medida particular que hacer parte de la línea de acción de Restitución, dentro del componente de Reparación Integral de la Política Pública para la Atención y Reparación Integral a las Víctimas del Conflicto Armado. El componente de reparación integral tiene el objetivo de reparar a las víctimas de manera integral, diferenciada, transformadora y efectiva, por el daño sufrido como consecuencia de las violaciones a los derechos humanos e infracciones al Derecho Internacional Humanitario –DIH- (UARIV, 2015, p. 3). Por su parte, la línea de Restitución busca el restablecimiento de la situación existente de las víctimas, antes de las violaciones a los derechos humanos y al DIH.

La reglamentación principal de la Política de Retornos y Reubicaciones se encuentra al interior de la Ley 1448 de 2011, del Decreto 4800 de 2011 y del Decreto 1084 de 2015. Dentro de esta normatividad se establece que el propósito de esta política es garantizar la atención integral a las víctimas de desplazamiento forzado que deciden voluntariamente retornar o reubicarse, con condiciones de seguridad adecuadas, y que a partir de esquemas especiales de acompañamiento se les pueda garantizar su goce efectivo de derechos (Ley 1448 de 2011, art. 66). Los encargados de adelantar este proceso son las distintas entidades pertenecientes al

[58] "El retorno es el proceso mediante el cual la persona o el hogar víctima de desplazamiento forzado deciden regresar al sitio del cual fueron desplazados con el fin de asentarse indefinidamente" (Decreto 4800 de 2011, art. 71).

[59] "La reubicación es el proceso mediante el cual la persona o el hogar víctima de desplazamiento forzado deciden asentarse en un lugar distinto del que se vieron forzados a salir" (Decreto 4800 de 2011, art. 72).

Sistema Nacional de Atención y Reparación a las Víctimas (SNARIV), en liderazgo de la UARIV[60].

El acompañamiento que se brinde por parte del estado para llevar a cabo procesos de retorno y reubicación debe estar dentro del marco de los principios de seguridad, voluntariedad y dignidad (Decreto 4800 de 2011). Con esto, se espera evitar la vulneración de los Derechos Humanos e infracciones al DIH, que la población que decida voluntariamente y con pleno conocimiento de las condiciones del lugar de destino, y brindar acceso a planes, programas y proyectos orientados al goce efectivo de derechos.

Para la ejecución de los procesos de retorno y reubicación, la UARIV coordina el diseño e implementación de planes de retorno y reubicación dentro de los Comités Territoriales de Justicia Transicional, los cuales son la herramienta para el diagnóstico, definición de responsabilidades, cronograma y seguimiento a los procesos (art. 78). Estos planes garantizan de manera gradual el goce efectivo de derechos de la población beneficiaria, priorizando los derechos de atención básica en salud, educación, alimentación, identificación, reunificación familiar, orientación ocupacional, vivienda y atención psicosocial (art. 75).

La elaboración de los planes y su implementación se realiza en el marco del Protocolo de Retorno y Reubicación (UARIV, 2017a), el cual constituye el principal documento de diseño de política pública. Dentro de este documento se establecen todos los lineamientos y requisitos que deben tener los planes de retorno y reubicación, se describen las fases para el acompañamiento estatal a estos procesos, y se describe cinco distintas rutas de acompañamiento dependiendo si se trata de procesos de retorno y reubicación para sujetos de reparación colectiva; procesos individuales y masivos; procesos dentro de territorios colectivos; procesos de emergencia y corto plazo; y procesos de reubicación en sitio de recepción.

[60] Con los derechos mínimos de: "… identificación a cargo de la Registraduría Nacional del Estado Civil, salud a cargo del Ministerio de la Protección Social, educación a cargo del Ministerio de Educación Nacional, alimentación y reunificación familiar a cargo del Instituto Colombiano de Bienestar Familiar, vivienda digna a cargo del Ministerio de Ambiente, Vivienda y Desarrollo Territorial cuando se trate de vivienda urbana, y a cargo del Ministerio de Agricultura y Desarrollo Rural cuando se trate de vivienda rural y orientación ocupacional a cargo del Servicio Nacional de Aprendizaje" (Ley 1448 de 2011, art. 66).

6.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Retornos y Reubicaciones

6.2.1 Reglas de alcance y resultados potenciales

Dentro de los elementos que se presentan a modo de justificación de la realización de una evaluación estratégica se encuentran: la meta del Plan Nacional de Desarrollo (Estrategia 6.5.) la cual planteó que para el 2018 alcanzar un total de 160.000 hogares acompañados en procesos de retorno y reubicación; una inversión anual de 9.1 billones de pesos para el 2015 por parte de las entidades del SNARIV; la implementación del punto 5 del Acuerdo Final entre el gobierno nacional y las FARC; la realización de un análisis holístico en conjunto con las evaluaciones concomitantes al Programa Familias en Su Tierra del Departamento para la Prosperidad Social –DPS- (en donde se atendieron a más de 38.000 hogares), y al Programa de Reparación Colectiva de la UARIV (en donde el retorno o reubicación han hecho parte de los planes de reparación colectiva); y el planteamiento de una estrategia para el fortalecimiento y ejecución de la Política de Retornos y Reubicaciones (DSEPP-DNP, 2016a, p. 3).

En cuanto a la finalidad de esta evaluación, se estableció que con esta se brindaría insumos al gobierno nacional para “evidencia frente al desempeño de la política y a la falta de información evidenciada frente a los procesos no acompañados institucionalmente, al igual que la identificación de los cuellos de botella y factores de éxito en la verificación de los principios de seguridad, dignidad y voluntariedad necesarios para la sostenibilidad de la política” (p. 4). Con esto, se observa que la evaluación está dirigida hacia un uso formativo, y además de desarrollo de capacidad.

El objeto del proceso de contratación de la consultoría fue realizar una evaluación de operaciones y resultados para “… identificar las fortalezas y debilidades de las acciones de atención a las víctimas de desplazamiento forzado, así como la sostenibilidad de los procesos de retorno y reubicación” (DNP, 2016a, p. 1), con un plazo de ejecución del contrato de nueve meses.

El alcance de la consultoría se dividió en los siguientes componentes: 1) evaluación de operaciones, que buscó la identificación de fortalezas, debilidades y vacíos en la elaboración y ejecución de los planes de retorno y reubicación; 2) Evaluación de resultados, el cual se dirigió a identificar las condiciones en las que vive la población beneficiaria de la política, en cuanto a goce efectivo de derechos y efectividad de los planes de retorno y reubicación; 3) Conclusiones y recomendaciones, el cual se direccionó hacia la formulación de recomendaciones y sugerencias para “hacer la política más eficiente” (DSEPP-DNP, 2016a, p. 60), y hacia la obtención de evidencia acerca de aspectos de la política que debieron ser ajustados para mejorar la operación del programa y permitir un mayor logro en los resultados. Puntualmente

se manifiesta que las recomendaciones de esta evaluación se formularán para hacer uso de estas en los procesos de toma de decisiones (p. 72).

Los anteriores elementos direccionaron adecuadamente los estándares de fundamentación y finalidad de la evaluación, de objeto y alcance de la evaluación. Así mismo, los objetivos específicos son aclarados en el Pliego de Condiciones Definitivo (p. 4) del proceso de contratación de la consultoría, en estos, se establece claramente el propósito de la evaluación, lo que satisface el estándar de objetivos específicos de la evaluación.

Además del abordaje de los problemas de falta de información, dentro de lo que se identificó como la finalidad de la evaluación, dentro de los documentos de diseño de la evaluación se encuentra menciones al Auto 202 de 2015 de la Corte Constitucional, en donde se consideró que no se cuenta con información detallada sobre el universo de casos de retornos y reubicaciones cubiertos por las políticas del gobierno, por lo que, la evaluación se consideró indispensable para definir criterios de selección de casos, indicadores, metodologías de medición y valoración (p. 3). La evaluación contratada al potenciar las capacidades de evaluación de los implementadores de la política, y del Estado en general, direccionó adecuadamente el estándar de desarrollo de capacidad.

Dentro de los lineamientos y requerimientos para la realización del PTIR, que se encuentran en el pliego de condiciones definitivo, se incluye la aclaración y descripción de la metodología que se utilizará para la identificación y priorización de las recomendaciones de la evaluación; la relación entre las recomendaciones y los objetivos del programa; la relación entre la evidencia aportada por la evaluación y las recomendaciones realizadas; las razones y enfoques utilizados para sustentar las recomendaciones; y la identificación de actores clave a quienes afectó la recomendación (p. 72). Al verificar estos elementos en el PTIR realizado por la consultoría, se encuentra que, en efecto, los componentes del estándar de oportunidad, pertinencia y utilización que corresponden a reglas de alcance, fueron adecuadamente abordados por el proceso de evaluación.

6.2.2 Reglas de frontera y participantes

Los participantes dentro del Comité Técnico de Seguimiento fueron el DNP (DSEPP y Grupo de Proyectos Especiales), la UARIV y el DPS, representados con ocho funcionarios en total (Econometría y SEI, 2017a, p. 3). La Procuraduría General de la Nación y la Defensoría del Pueblo, como integrantes de la Comisión de Seguimiento y Monitoreo a la Implementación de la Ley 1448 de 2011, participaron al interior de la mesa de expertos que ayudó a la formulación de las recomendaciones de la evaluación.

Dentro de los actores con los cuales la UARIV debe coordinarse para la implementación de las recomendaciones resultantes de la evaluación (según la información del PTIR), y que estuvieron ausentes dentro del Comité de Seguimiento (y en general de la participación dentro del proceso de evaluación) se destacan: la Unidad de Restitución de Tierras, el Ministerio del Interior; Congreso de la República; Presidencia de la República; Ministerio de Hacienda y Crédito Público; y Departamento Administrativo para la Función Pública). Con esto, sólo hay un cumplimiento parcial de los estándares de participación de los actores relevantes, consideración de evaluación conjunta, y de enfoque asociativo.

Con respecto al equipo de evaluación, los requisitos de este se aclararon en varios documentos, incluyendo el Aviso de la convocatoria para la contratación de la consultoría (DNP, 2016a). Dentro de este documento se aclaran que la modalidad de selección será a través de concurso de méritos abierto. Para este último, se aclararon requisitos habilitantes jurídicos; el calendario de presentación de las propuestas (a más tardar el 20 de septiembre de 2016); los requisitos habilitantes técnicos; requisitos habilitantes financieros; y los criterios de calificación de las propuestas.

Estos mismos requisitos son aclarados en el pliego de condiciones definitivo, y se adiciona la aclaración del perfil del equipo de la evaluación requerido. Entre estos se encuentran: director del proyecto, que debió tener nivel educativo mínimo de maestría, con tres años de experiencia en la dirección de investigaciones o evaluaciones de programas del sector de inclusión social; Experto en población desplazada con mínimo maestría y experiencia de dos años en formulación, análisis o evaluación de proyectos relacionados; analista cuantitativo; y analista cualitativo (DSEPP-DNP, 2016a, pp. 19-20). Dependiendo del nivel educativo o de los años de experiencia que se tuviesen adicionales a los requerimientos planteados, se asignaba un mayor puntaje a la propuesta. Todo esto, satisfizo el estándar de equipo de evaluación.

Por último, con respecto a los recursos, el DNP pagó el valor del contrato, el cual fue de $830.166.392 (DNP, 2016b, p. 4). Esto se pagó en tres pagos de $166.033.280, un cuarto pago de $207.541.600, y un quinto pago de $124.524.958, todos a partir de la recepción a satisfacción de los cinco productos de la consultoría. Por su parte, el consultor presentó su propuesta económica, la cual fue aceptada y garantizó la suficiencia de recursos. Esto, más los puntos anteriores garantizaron la satisfacción del estándar de recursos.

6.2.3 Reglas de elección y acciones

En cuanto al análisis de evaluabilidad se encontró que la labor principal en este punto, correspondiente a la realización de una adecuada definición de la intervención, quedó a cargo de los proponentes participantes en el concurso de méritos abierto. Esto se refleja dentro de los requisitos incluidos en el pliego:

> “Los proponentes deberán identificar y analizar la cadena de valor de la intervención, de tal forma que se presente de forma clara y detallada los insumos, procesos, productos y resultados asociados a la política pública. Se espera que más allá de la descripción, el proponente sea capaz de hacer un análisis de cada eslabón de la cadena de valor, así como una articulación entre ellos” (DSEPP-DNP, 2016a, p. 28).

En esta reconstrucción de la cadena de valor dentro de sus propuestas, los proponentes debieron incluir un análisis de “cuellos de botella” dentro de sus eslabones. Así mismo, debieron incluir propuesta de las variables, indicadores, categorías de análisis y de fuentes de información primaria y secundaria para responder a las preguntas que se incluyeron en el pliego y en el aviso de convocatoria (DNP, 2016a, p. 29). En este sentido, a pesar de los problemas que puedan identificarse con respecto a las falencias del Comité de Diseño en el direccionamiento de elementos fundamentales de la evaluación, y en cuanto a la asignación presupuestaria a un contrato de evaluación que no tiene aún análisis de evaluabilidad, los componentes de este estándar fueron satisfechos del modo descrito.

La Política de Retornos y Reubicaciones requiere de la coordinación de las distintas entidades que componen el SNARIV para su efectiva ejecución, dentro de estas se encuentran entidades de distintos sectores y de los tres niveles de gobierno, por lo que se debe tener consideración de los planes y las políticas nacionales y locales, con los cuales los planes de retorno y reubicación deben estar armonizados.

La evaluación parte del marco normativo que incluye la Ley 1448 de 2011, el Decreto 4800 de 2011, el Decreto 1084 de 2015, los Principios Rectores de los Desplazamientos Internos -Principios Deng- (ONU, 1998), la Sentencia T-025 de 2004, Adicionalmente, ubica el proceso de retornos y reubicaciones dentro de la ruta general de atención, asistencia y reparación integral para desplazados (Econometría y SEI, 2017a, pp. 30-31). Esto significa que el estándar de coordinación y alineamiento fue satisfecho en el proceso de evaluación.

En la acción correspondiente a la selección de la metodología sucede similar al análisis de evaluabilidad, en este caso se brindó gran discrecionalidad al proponente y, por ende, al equipo evaluador de identificar y seleccionar (por lo menos en una primera instancia) un conjunto considerable de elementos de la metodología. Dentro del pliego de condiciones definitivo se aclaró que los proponentes debieron especificar los siguientes puntos de la metodología dentro de sus propuestas (DSEPP-DNP, 2016a, pp. 27-33):

- Metodologías cualitativa y cuantitativa para la evaluación de operaciones.
- Metodologías cualitativa y cuantitativa para la evaluación de resultados.
- Descripción y justificación de cada una de las cuatro metodologías

Una vez el consultor fue seleccionado, el primer producto consistió en un diseño metodológico, el cual requirió el perfeccionamiento de múltiples aspectos de la propuesta metodológica, y la adición de otros nuevos, entre los cuales se destaca:

- Especificación del enfoque conceptual, variables de análisis, indicadores, y método de evaluación y valoración de cada variable (p. 65).
- Proceso de levantamiento y análisis de información (p. 65).
- Descripción y justificación de instrumentos de recolección de información (p. 65).
- Matriz de consistencia de la evaluación que relacione los objetivos, las preguntas orientadoras, las variables, los indicadores y las preguntas de los instrumentos de recolección de datos (p. 67).
- Diseño y justificación de los muestreos cualitativo y cuantitativo (p. 67).
- Esquema general de codificación para el análisis de información cualitativa (p. 68).
- Esquema metodológico para la formulación de recomendaciones (p. 68).

Con esto, todos los componentes del estándar de selección de enfoque y método fueron completamente satisfechos. Así mismo, se definen un total de dieciséis preguntas para el componente de evaluación de operaciones, nueve preguntas para el componente de evaluación de resultados; y cuatro preguntas para el componente de conclusiones y recomendaciones, para un total de veintinueve preguntas plasmadas al interior del pliego de condiciones definitivo. Dentro de estas preguntas orientadoras (pp. 59-60), los criterios de evaluación que son abordados son eficiencia, pertinencia, eficacia, y sostenibilidad. Todo esto direccionó adecuadamente los estándares de cuestiones a responder y selección y aplicación de criterios de evaluación.

Con respecto al conjunto de acciones que se deben realizan, una vez se tiene el diseño de la evaluación y esta se comienza a ejecutar, en primer lugar, se destaca la adecuada organización de las respuestas a las preguntas orientadoras al interior del informe final de la evaluación. Particularmente, en la primera parte del informe final, se cuenta con una matriz (Econometría y SEI, 2017a, p. 7) que relaciona la ubicación dentro del documento en el cual está la respuesta a la pregunta orientadora de evaluación, el objetivo específico de la evaluación del cual hace parte, y el objetivo del proceso de retornos y reubicaciones con el cual se relaciona la pregunta orientadora. No se encontraron preguntas sin responder. Gracias a lo anterior, los estándares de respuestas a preguntas evaluativas y de claridad del análisis fueron satisfechos completamente.

La entrega del informe final fue acompañada con un resumen ejecutivo, el cual tuvo lineamientos claros al interior del pliego de condiciones definitivo en cuanto a su claridad, facilidad de comprensión y síntesis (DSEPP-DNP, 2016a, p. 70). Este documento de resumen ejecutivo (Econometría y SEI, 2017b) incluyó una síntesis de la metodología, presentación de hallazgos de la evaluación de operaciones, presentación de los hallazgos de la evaluación de resultados. Las conclusiones y

resultados son presentados en la ficha técnica de la evaluación. Esto satisfizo el estándar de claridad y representatividad del resumen de la evaluación.

Frente al abordaje del contexto de la intervención por parte de la evaluación, además de ser requerido como un aspecto particular a abordar dentro del componente de evaluación de operaciones en el pliego de condiciones (DSEPP-DNP, 2016ª, p. 61), las variables del contexto de la intervención son abordadas por las preguntas orientadoras de la evaluación, y son determinadas a partir de los principios que guían la implementación de la política (seguridad, voluntariedad y dignidad). De esta manera, el estándar de contexto de la intervención quedó satisfecho dentro del informe final.

En el informe final que se puede acceder por internet no se encontró evidencia de una reconstrucción exhaustiva de la lógica de la intervención. Dentro de este documento se aclara que el acompañamiento estatal -como elemento central de la Política de Retornos y Reubicaciones- no se ha elaborado como una política pública sino como un proceso dentro de toda la Política de Prevención, Protección, Atención, Asistencia y Reparación Integral a las Víctimas (Econometría y SEI, 2017a, p. 25).

El argumento central para sostener que no se trata aun de una política pública como tal, se basa en que el principal documento de diseño de la política, el Protocolo para el Acompañamiento a Retornos y Reubicaciones en el Marco de la Reparación Integral a las Víctimas, genera una confusión con respecto al alcance de la política, el cual, al ser tan amplio “incluso implicaría fusionar el proceso de retornos y reubicaciones con toda la política de atención y reparación a las víctimas” (p. 29).

En todo caso, se menciona un análisis del protocolo como documento de diseño de política, pero al no incluirse los anexos dentro del documento destinado a ser difundido, en los cuales es donde se menciona que se presenta este análisis (p. 30), es imposible realizarle una valoración a la presentación de la lógica de la intervención. Esto significa que el estándar de lógica de la intervención sólo puede considerarse como parcialmente satisfecho.

En el informe final se realiza una adecuada presentación de la metodología de evaluación, la cual incluye la descripción de las herramientas metodológicas de evaluación (p. 5), del proceso de evaluación (p. 6), el conjunto de factores críticos como herramienta conceptual de análisis (p. 11), las fuentes e instrumentos de recolección de información (p. 12), el tipo de análisis realizado para cada componente (p. 13; p. 16) y el diseño muestral y de trabajo de campo (p. 17).

En este último se aclara un total de 1250 hogares como muestra de estudio en todo el país, 20 grupos focales, 77 entrevistas a funcionarios territoriales, 31 entrevistas a funcionarios del nivel nacional, cuatro entrevistas etnográficas y seis estudios de caso. Finalmente, y siguiendo los lineamientos del pliego de condiciones, se

presenta la estrategia de triangulación de la información cuantitativa y cualitativa a lo largo del capítulo de presentación de la metodología, con lo que se satisfacen todos los componentes de los estándares de explicación de la metodología, y de validez y fiabilidad de las fuentes de información.

En cumplimiento con los requerimientos del pliego de condiciones (DSEPP-DNP, 2016a, p. 28), se realizan menciones acerca de los desafíos y límites que tiene la metodología de evaluación. Dentro de estas, están los problemas existentes con las bases de datos disponibles para determinar la muestra, en donde se encontró un 27,6% de registros sin información del RUV ni del SISBEN, en la información entregada por el Comité de Seguimiento, lo que no permitió localizar las unidades de estudio (Econometría y SEI, 2017a, p. 18), de la muestra seleccionada, no se pudo localizar a 894 registros, lo que significó un 56,2% de pérdida de la muestra; también problemas con respecto a la clasificación de los sujetos en los grupos focales (p. 21); y problemas en ubicar unidades de observación que se encontraban por fuera del país en la construcción de los estudios de caso (p. 21). De este modo, el estándar de mención de cambios y límites fue satisfecho, pero no se realiza mención de discrepancias entre miembros del equipo evaluador, por lo que este último estándar no se cumplió.

En cuanto al estándar de consulta y protección de implicados, se observa un abordaje adecuado de los actores relevantes a ser consultados, a nivel nacional fueron: “Unidad para la Atención y Reparación Integral a las Víctimas, Ministerio de Salud y Protección Social, Ministerio del Interior, Ministerio de Vivienda, Ciudad y Territorio, Unidad de Restitución de Tierras, Ministerio de Defensa Nacional, Cancillería, la Contraloría General de la República, la Procuraduría General de la Nación, Defensoría del Pueblo, ACNUR, el Consejo Noruego para Refugiados (CNR) y la Comisión de Seguimiento a la Política Pública sobre Desplazamiento Forzado (CSPPDF)” (DSEPP-DNP, 2016a, p. 32), mientras que a nivel territorial se consultó a 161 municipios con planes de retorno y restitución aprobados, y 79 con medidas implementadas.

De conformidad con la obligación del DNP de anonimizar la información producto de la consultoría, para salvaguardar el derecho a la intimidad, habeas datas y buen nombre (p. 6), en el informe final se evidencia que los datos de las entrevistas fueron presentados sin nombres propios, con esto se cumplió con el estándar de consulta y protección de implicados.

Por último, en el contrato de consultoría (DNP, 2016b, p. 7) se establece como fecha de entrega del informe final el 28 de abril de 2017, y del PTIR el 28 de mayo de 2017, sin embargo, al dirigirse a los productos, se evidencia que el informe final tiene fecha del 28 de agosto de 2017 y el PTIR del 11 de agosto de 2017. Esto significa que hubo un retraso de cuatro y tres meses respecto de la fecha en la cual el Comité de Seguimiento estableció como pertinente para contar con los productos. Esto perjudica la oportunidad y relevancia de la evaluación y hace que el estándar de evaluación dentro del plazo y presupuesto asignado no haya sido satisfecho.

6.2.4 Reglas de agregación y control

Dentro del pliego de condiciones definitivo se establecen las principales reglas de agregación. Dentro de estas se destacan que, el consultor debió entregar informes técnicos cuando le sean solicitados por el supervisor del contrato acerca del avance en cualquiera de los productos requeridos, en un término no mayor a ocho días hábiles (DSEPP-DNP, 2016a, p. 5); el requerimiento de póliza de seriedad de la oferta, en la que se garantizó el cumplimiento de esta por parte del consultor por el 10% del valor del presupuesto del contrato (DNP, 2016b, p. 13), la cual cubrió riesgos en cuanto a: cumplimiento del contrato, calidad del servicio, pago de salarios, prestaciones sociales e indemnizaciones laborales (p. 55); la cláusula penal compensatoria (p. 11), con una pena pecuniaria del 30% del valor del contrato en caso de incumplimiento parcial o total por parte del consultor; las multas que considera el contrato en caso de mora o incumplimiento parcial con un tope máximo del 20% del valor del contrato (p. 11); y la aprobación previa de todos los ajustes metodológicos por parte del supervisor del contrato (DSEPP-DNP, 2016a, p. 68).

Como en las demás evaluaciones la supervisión estuvo a cargo del Comité de Seguimiento, y por parte de la Secretaría Técnica del Comité de Seguimiento (DSEPP). Esta última también tuvo a cargo la consolidación, depuración y tratamiento de los comentarios realizados por parte de los miembros del Comité de Seguimiento a los productos entregados por parte del consultor (p. 56). Posteriormente, se tuvo estipulado:

> “Con base en los resultados de los comentarios y sus discusiones, el CONSULTOR realizará los ajustes solicitados por el Comité y enviará el informe definitivo en físico y en medio magnético al supervisor del contrato” (p. 56).

> “De acuerdo con los comentarios y sus discusiones, el CONSULTOR realizará los ajustes solicitados por el Comité y enviará el informe ajustado en medio electrónico con el fin de que dicho Comité verifique los ajustes. Una vez aprobado, el CONSULTOR enviará el producto en físico y en medio magnético al supervisor del contrato para proceder con el recibido a satisfacción para pago” (p. 73).

Como se observa, se estipula un control de calidad a lo largo de todo el proceso de evaluación a partir de diversos mecanismos: a la revisión de los productos por parte de la Secretaría Técnica y de los miembros del Comité de Seguimiento se le suma la revisión de las recomendaciones a la mesa de expertos (lo que se aborda más adelante). De esta manera, todos los componentes de los estándares de control de calidad e inclusión de observaciones de los implicados fueron satisfechos.

6.2.5 Reglas de posición

Ciertamente, en esta evaluación existieron normas a favor de la independencia del equipo evaluador, lo cual no existe en las evaluaciones anteriores. Al interior del contrato existe una cláusula a favor de la autonomía del evaluador (DNP, 2016b, p. 11), en esta se plantea que el consultor actuará con total autonomía técnica y administrativa en el cumplimiento de sus obligaciones contractuales. Por otra parte, el evaluador contó únicamente con un plazo de una semana para entregar el producto con los ajustes solicitados por parte del Comité de Seguimiento (p. 5), a riesgo de que el DNP activara las cláusulas de multas, mecanismos de prevención del riesgo, o la penal compensatoria. Con respecto a la naturaleza de los comentarios realizados por el Comité de Seguimiento de los conceptos técnicos de los productos se destacan los siguientes:

> "En general, no se encontró una respuesta en profundidad de las preguntas orientadoras. Es importante garantizar que las preguntas sean respondidas y que la respuesta sea a fondo" (DSEPP-DNP, 2017a, p. 3).
>
> "Suelen mencionarse en el documento, sin distinción, afirmaciones que corresponden a opiniones de los entrevistados y opiniones de la firma. Es necesario, que cuando se quieran presentar opiniones de víctimas o funcionario se aclare que son opiniones. Y que las posiciones que adopte la firma se sustenten" (p. 3).
>
> "Si no se presentan argumentos fuertes con base en el testimonio y en la evidencia encontrada cualquiera que sea el instrumento aplicado o la revisión de la información secundaria las afirmaciones pierden total validez" (p. 4).
>
> "Se sugiere llevar todo el desarrollo de resultados al informe correspondiente y centrar este informe en las recomendaciones, simplemente mencionando a qué resultado corresponden y por qué contribuyen a solucionarlo" (2017b, p. 4).
>
> "Las recomendaciones tienen a estar muy poco desarrolladas y más que centrarse en ajustes de fondo, muchas de ellas se refieren al ajuste de instrumentos o herramientas. No pareciera que la implementación de estas recomendaciones fuera a solucionar las problemáticas graves de la política (…) En este sentido, se sugiere por una parte ser más detallados en la forma de implementar las recomendaciones propuestas, y por otra parte centrarse en los temas de fondo y no en los instrumentos" (P. 4).

Los anteriores fueron los comentarios más centrales y relevantes que el Comité de Seguimiento realizó al producto del informe final y al PTIR. Tal como se puede observar, a pesar de que la estructura institucional de reglas de control restringe la discrecionalidad del evaluador, y que estos comentarios fueron aceptados sin mayor discusión por parte del equipo evaluador (tal como es descrito en los conceptos técnicos), la naturaleza de los comentarios evidencia que no se utilizó las

correcciones de estos productos como medio de restricción de la independencia del evaluador. Esto significó que tanto el estándar de proceso evaluativo libre y abierto, como el de independencia del evaluador, fueran satisfechos en este proceso de evaluación.

6.2.6 Reglas de información

En el portal del SECOP se puede encontrar la información respectiva a este contrato de consultoría. Al interior de este portal se encuentra el documento con los estudios previos de la contratación, el proyecto de pliego de condiciones, y el pliego de condiciones definitivo. En este último se aclaran los términos de referencia relativos al alcance de la evaluación, sus objetivos, lineamientos metodológicos generales, el monto del contrato, y el cronograma que inició con la adjudicación del contrato. Por lo que el estándar de documentación acerca de la finalidad y la expectativa de la evaluación fue satisfecho a cabalidad.

Con respecto a la difusión y a la claridad de los productos elaborados, se observa un cumplimiento de los lineamientos del pliego de condiciones relacionados con la forma y el contenido de cada uno de los productos. El informe final estuvo acompañado del resumen ejecutivo, de la ficha técnica, el PTIR, y dos presentaciones: una extendida con la explicación exhaustiva de la metodología, resultados y recomendaciones; y una simple con una breve presentación de la metodología, los resultados y las recomendaciones más relevantes, cuya priorización fue objeto de deliberación en conjunto con el Comité de Seguimiento (DSEPP-DNP, 2016a, p. 72). También se requirió un video con los mismos elementos de la presentación simple, con una duración entre 5 y ¿? minutos.

Además de la difusión regular que se realiza a los productos dentro del Comité de Seguimiento, y en el portal de SINERGIA, los últimos dos productos de la consultoría de esta evaluación fueron difundidos ante los directivos del DNP, la UARIV y el DPS (p. 5). De tal manera, los elementos correspondientes a reglas de información del estándar de informe de evaluación y el estándar de difusión fueron satisfechos completamente.

6.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la Política de Retornos y Reubicaciones

Teniendo en consideración la similitud conceptual que existe entre los estándares correspondientes a reglas de frontera, y el lineamiento de hallazgo e involucramiento de los usuarios previstos en el proceso de evaluación, como componente de la coordinación, se debe considera que a pesar de que se cuenta con los principales actores encargados de la formulación de la Política de Retornos

y Reubicaciones a nivel nacional, muchos actores que tienen competencias fundamentales para la adecuada implementación de las recomendaciones de la evaluación, no fueron incluidos en el proceso de evaluación más allá de su consulta como fuentes de información (dentro de los cuales, el Ministerio del Interior es probable que sea el de mayor importancia, considerando el papel que juega la articulación entre niveles de gobierno para esta política).

La inclusión de todos los actores involucrados en la política pública tiene, entre otras ventajas, la capacidad de aumentar la relevancia, oportunidad y apropiación de los hallazgos de la evaluación. Como se observará en la medición acerca de la efectividad del proceso de evaluación, en estos momentos, y desde que se tienen los resultados de la evaluación, la política pública se encuentra en un periodo de reformulación a través de distintos escenarios de coordinación interinstitucional. En este sentido, al dejar de incluir un número considerable de actores dentro del Comité de seguimiento, se perdió la capacidad de llegar a acuerdos y aprendizajes que actualmente se siguen intentando construir.

En cuanto a la relación que se sostuvo dentro del equipo de trabajo conjunto entre DNP, consultor y usuarios previstos, al preguntar a los funcionarios de esta política al respecto, se resaltó un adecuado grado de reactividad del DNP y el equipo evaluador con respecto de la preocupaciones y demandas de los usuarios previstos:

> “La interacción fue muy buena en términos digamos del acompañamiento que brinda DNP, el liderazgo que da al proceso de la evaluación es importante y es muy bueno, en qué sentido, por ejemplo para la determinación de los universos, si yo tengo personas que abordar, se tiene en cuenta a la entidad a la hora de identificar cuáles son esos actores fundamentales (…) efectivamente la instancia son esos comités técnicos que lidera el DNP y donde se identifica cuáles son las necesidades de información y demás” (Asesora para la formulación de política pública, Subdirección General-UARIV).

Como se observa, la percepción positiva acerca de la coordinación enfoca la reactividad en las discusiones metodológicas que se dieron en el equipo de trabajo conjunto. Esto pudo ser confirmado en varios modos: por un lado, en la metodología para la caracterización del conjunto de factores críticos que guio el análisis de procesos, se entrevistó a los usuarios previstos, y se realizó un taller con el Comité de Seguimiento (Econometría y SEI, 2017a, p. 10).

También, en la descripción de la metodología para la identificación y priorización de recomendaciones (Econometría y SEI, 2017c, p. 12), en donde se aclara que se parte de la identificación de problemas, barreras y dificultades a partir de las conclusiones y los hallazgos de la evaluación, y la estructura en un árbol de problemas; posteriormente se realizó una identificación de acciones ya implementadas para el abordaje de estos problemas; luego se procedió a identificar las medidas de ajuste estructural y los cambios requeridos para su implementación; a continuación se realizó una descripción y justificación de las recomendaciones, y

se las organizó en una matriz; la siguiente fase fue validar esas recomendaciones en una mesa de expertos; finalmente, se aclara que se realizó la discusión y el ajuste de las recomendaciones con el Comité de Seguimiento.

Así mismo, la reactividad del equipo evaluador en la adaptación de la metodología a lo largo de la ejecución de la evaluación se observa en la inclusión de los comentarios que el Comité de Seguimiento realizó al producto N°1, correspondiente al informe metodológico por parte del equipo evaluador. Dentro de los comentarios más sobresalientes, se encuentran (DSEPP-DNP, 2017c, pp. 2-5):

- Dejar abierta la posibilidad de identificar nuevos factores críticos de éxito con las preguntas incluidas en los instrumentos de recolección de datos.
- Establecer el tratamiento de los datos en la triangulación de estos, cuando los datos cualitativos vayan en sentido opuesto a los cuantitativos.
- Diferenciar los análisis del “deber ser”, “ser” y “viabilidad del diseño”.
- Centrar el análisis de los componentes principales con variables (indicadores) continuas y no categóricas (ordinales o nominales).
- Incluir en el componente de operaciones el análisis de implementación de la política a nivel nacional.
- Incluir en el análisis de ejecución presupuestal la inversión a nivel sectorial.

En el informe final, también, se aclara que diversos puntos de la metodología se acordaron en conjunto con el Comité de Seguimiento (Econometría y SEI, 2017a, p. 30), entre los que se encuentran las temáticas y los municipios con los que se elaboraron los estudios de caso, y el muestreo cualitativo. Así mismo, el Comité de Seguimiento tuvo un papel determinante en la conceptualización para la clasificación de los hogares, la construcción de indicadores y formulas (p. 13), y estableció los componentes con los cuales se midió a los hogares víctimas de desplazamiento que han sido acompañados (p. 139). Dentro de estos componentes se encuentran los hogares que, han recibido algún Esquema Especial de Acompañamiento a través de la ruta individual; los connacionales que han sido acompañados; los que cuentan con un acta de voluntariedad; los que están en el marco de procesos de cofinanciación de las entidades territoriales; los que cuantan con subsidio de vivienda; los que están en el Registro de Tierras Despojadas de la Unidad de Restitución de Tierras; los que cuentan con sentencia de restitución de tierras; o cuentan con un proyecto productivo.

También se evidencia un involucramiento de los usuarios previstos en la interpretación de los hallazgos y en la construcción de las recomendaciones. Adicional a la inclusión de los comentarios que fueron presentados en el análisis del cumplimiento de las reglas de posición, para la construcción de las recomendaciones, se partió de una propuesta realizada por parte del equipo evaluador, luego de esta propuesta se puso en consideración de una mesa de expertos, de esta mesa surgió una nueva propuesta, la cual fue finalmente puesta a consideración del Comité de Seguimiento. Al respecto -aparte del lineamiento del

contrato y del pliego de condiciones definitivo, acerca de la modificación del PRITR a partir de los comentarios del Comité de Seguimiento- se encontró confirmación de la realización de esto dentro del documento PTIR:

> "Documentación, discusión y ajuste con el Comité Técnico de Seguimiento (…) Corresponde a la elaboración del documento PTIR en su versión preliminar, así como la discusión, validación y ajuste con el Comité Técnico de Seguimiento (CTS). A partir de los comentarios y ajustes sugeridos por el CTS, la consultoría elabora la versión final del Plan de Transferencia e Implementación de las Recomendaciones" (Econometría y SEI, 2017c, p. 16).

No se encontró evidencia de un plan conjunto de seguimiento a la implementación de los hallazgos de la evaluación. Sin embargo, dentro de los componentes de la coordinación de los participantes en esta evaluación se destaca que existen acciones que reflejan una búsqueda proactiva del uso, lo cual no se había visto en las evaluaciones anteriores. Si bien no ha existido una búsqueda constante de oportunidades por parte del evaluador, para el reforzamiento de la relevancia de los hallazgos, se encontró la siguiente información:

> "… otro tipo de relacionamiento es ya cuando están los resultados, es el tema de la disposición que tuvo por ejemplo tanto DNP como la misma firma evaluadora, que es este caso fue econometría, de poder socializar con la institución en las diferentes instancias, acá hicimos dos o tres procesos de socialización de la evaluación de tal manera que efectivamente todo lo que había consignado en el documento realmente era entendido tanto por directivos, como por funcionarios de la entidad para apoyar todo ese proceso de redefinición que venía posteriormente. Es poder dar más información, más argumentos, más contexto, porque una cosa es lo que dice un documento y otra cosa es cuando lo exponen quienes lo han construido, entonces aclarar dudas y demás" (Asesora para la formulación de política pública, Subdirección General-UARIV).

Como se observa, las acciones referidas a la difusión de los resultados de este proceso de evaluación fueron mucho más allá que en los procesos de evaluación anteriores. En este caso se aumentó la relevancia de los hallazgos realizando socializaciones a cargo del mismo evaluador, en diversos niveles de la UARIV. Como se observa, el hecho que sea el evaluador y no el cliente de la evaluación, el encargado de realizar esta difusión de los hallazgos tiene un importante componente en la defensa de estos últimos.

6.4 Análisis de la efectividad de la evaluación de operaciones y resultados de la Política de Retornos y Reubicaciones

Esta evaluación produjo únicamente tres recomendaciones. A pesar de su escaso número, estas recomendaciones tuvieron un grado considerable de complejidad gracias a las acciones y medios de implementación requeridos para su completa aplicación. En primer lugar, se recomendó una redefinición conceptual completa de todo el proceso de acompañamiento estatal a los procesos de retorno y reubicación (Econometría y SEI, 2017c, p. 24).

Esta recomendación requirió: cambios normativos para la delimitación del alcance de los procesos y esquemas de acompañamiento, para lo cual es necesario un análisis y ajuste de los decretos reglamentarios vigentes en la materia; requirió de cambios en los instrumentos de implementación de los procesos, a través de cambios a los protocolos, manuales y guías; y requirió cambios institucionales orientados hacia el fortalecimiento del grupo de la UARIV encargado de retornos y reubicaciones; y el fortalecimiento de los retornos y reubicaciones como tema dentro de las direcciones territoriales de la UARIV. Tal como se observa, la recomendación plantea una reformulación estructural muy profunda de todo el proceso de política pública. Al respecto de su implementación se comenta:

> "Tomamos mucho del resultado de la evaluación sobre todo a la hora de justificar el cambio, porque se hizo un estudio y el estudio arrojo unos resultados. Por ejemplo aquí, definición conceptual del proceso de acompañamiento, aún estamos en este proceso (…) Ha tomado tiempo porque es un dialogo interno y externo, ha sido un dialogo interno con el DPS, ellos tiene un programa que se llama Familias en su tierra, para retornados y reubicados y tienen una opinión frente al nuevo procedimiento, el DNP no por el lado de econometría, sino por otra oficina también se ha vinculado al proceso de cambio que estamos haciendo" (Funcionario Grupo de Retornos y Reubicaciones, UARIV).

> "Si hablamos de la redefinición conceptual del proceso de acompañamiento, efectivamente en eso se está trabajando (…) cuando nosotros estamos trabajando en esa redefinición conceptual del proceso de acompañamiento, pues obviamente no solo tomamos los resultados de la evaluación sino tomamos en cuenta todo lo que nos han dicho las 'ias' [Ministerio Público], todo lo que nos ha dicho la Comisión de Seguimiento a la Política Pública, todo lo que nos ha dicho la Corte Constitucional y todo lo que internamente también se va encontrando, porque los mismos funcionarios documentan sus procesos y los mismos funcionarios o servidores públicos van identificando con su experiencia, las buenas prácticas de van documentando y se van identificando que cosas deben definitivamente mejorar (…) El tema de las recomendación asociada al fortalecimiento del grupo de retornos y

> reubicaciones dentro de la entidad, de alguna manera la recomendación particular de econometría era que cambiara de ubicación en la estructura organizacional el Grupo de Retornos y Reubicaciones, entonces uno dice no, esa es una de las recomendaciones que no se ha adoptado, pero no obstante, para el desarrollo y la ejecución operativa del programa ajustado, el grupo si debe reorganizarse (…) Los cambios normativos, claro resulta que si yo redefino pues si uno tiene como una guía en términos generales, y siempre esa guía en Colombia debe tener algo normativo o judicial o legal que lo soporte y entonces digamos que como estamos en el proceso de redefinición, pues hasta que no terminemos ese proceso pues esto no quedara normado a través de decretos o a través del ajuste al Decreto 1084 o a través de resoluciones internas, e incluso no solo las resoluciones internas sino justamente en el marco de los ajuste institucionales, también está toda la actualización de los mapas de procesos, procesos y procedimientos del sistema de gestión y calidad, todo debe ajustarse de acuerdo con esa redefinición" (Asesora para la formulación de política pública, Subdirección General-UARIV).

La redefinición conceptual del alcance de los procesos de retorno y reubicación es una recomendación con un grado considerable de abstracción, lo que implica, por un lado, desarrollar todo un proceso de coordinación interinstitucional entre los sectores de gobierno implicados, teniendo en consideración distintos insumos a parte de la evaluación realizada. Por otra parte, y con un mayor grado de complejidad, implica cambios normativos que requieren de la participación de entidades como Ministerio de Hacienda, Presidencia de la República, o incluso, del Congreso de la República.

Las falencias en el hallazgo e involucramiento de todos los usuarios potenciales, dentro del proceso de evaluación, impidió que la misma evaluación sirviera al proceso de coordinación institucional requerido para la redefinición conceptual, el cual, a año y medio después tener los resultados de la evaluación se observa muy lejano a tener resultados en esta dirección.

Por otro lado, lo ideal sería que en el mismo proceso de evaluación se tomaran en cuenta todos los insumos que se requirieran para hacer las recomendaciones que fuesen pertinentes. El hecho de que la evaluación sea tomada en cuenta solo como un referente entre muchos otros para los cambios que se observan como necesarios, denota fallas en su capacidad para dar cuenta de todo el contexto de la intervención evaluada.

Todo esto significa que a pesar de que los funcionarios sostengan que "se está trabajando en ello", la redefinición conceptual de los procesos de retorno y reubicación implica una transformación tan profunda de la política, que prácticamente implica iniciar un nuevo proceso de formulación de política pública. Ninguno de los "medios de implementación" de la recomendación se ha visto materializado a 2019, y con el tiempo, estos resultados van perdiendo cada vez más relevancia. En estos términos, la recomendación no se ha implementado, y es muy

poco probable que se llegue a implementar en los términos sugeridos por la evaluación, y no refleja efectividad del proceso de evaluación.

La segunda recomendación consistió en generar incentivos para el involucramiento local en soluciones que sean duraderas para la población beneficiaria (Econometría y SEI, 2017c, p. 24). Esta recomendación requirió como medios de implementación: cambios normativos y de política dentro del documento CONPES y con proyecto de reforma a la Ley 1448 de 2011; de cambios operativos en cuanto a la asistencia técnica en formulación de proyectos de inclusión social y productiva; ajustes normativos para la articulación y gestión de recursos; y ajustes en instrumentos de gestión de recursos, particularmente en cuanto a la revisión y adecuación a los convenios de cooperación internacional. Con respecto a su estado de implementación se comentó:

> "… está de incentivos para el involucramiento local y soluciones duraderas, entonces aquí está el marco de soluciones duraderas de la ONU, que es un documento que sale en el 2010, entonces esto no lo cogimos por Econometría, lo cogemos por el marco de soluciones, y comienzan a salir una serie de argumentos, por ejemplo este marco habla de que hay un concepto de reubicación que cuando la persona ya lleva mucho tiempo de arraigo le llaman integración local, entonces de ahí se ha venido cogiendo ese concepto y en el nuevo protocolo ahora se está empezando a plantear una definición de integración local y empezado a hacer unos pilotos en algunas zonas del país, el tema de integración local, pero lo estamos sacando del marco de soluciones duraderas que es un marco para el desplazamiento forzado y que la misma Corte en algunos autos también ha comenzado a ver. El gran actor aquí es la Corte, ellos hacen un estudio que lo que hace es que complementa mucho más lo que la Corte ya dijo" (Funcionario Grupo de Retornos y Reubicaciones, UARIV).

Como se puede observar, se han implementado acciones orientadas hacia las soluciones duraderas en los territorios. Sin embargo, esto no se da gracias a la recomendación del equipo evaluador en conjunto con la mesa de expertos, sino que se da a causa de los desarrollos que la ONU y la Corte Constitucional hacen en la materia, previo a la evaluación. La recomendación no significó una transformación en el modelo mental de los usuarios, justifica decisiones previas, y no refleja efectividad del proceso de evaluación.

Por último, se recomendó realizar una reingeniería de los flujos de conocimiento (Econometría y SEI, 2017c, p. 24), que implicaba realizar una serie de cambios en los instrumentos de manejo de la información. Entre estos se encuentran: diagnóstico y articulación de bases de datos existentes; diagnóstico de los medios y canales de comunicación con las víctimas; acceso a recursos tecnológicos a nivel local; sistema de cruce y consulta de información; construir un sistema de alertas

sobre seguridad; e incorporar la verificación de voluntariedad y dignidad en las encuestas de caracterización de beneficiarios. Con respecto a su uso se comentó:

> "En eso también se viene trabajando justamente en el proceso de definición del mejoramiento de procesos (…) más que sea un problema de sistematización de la información sobre este ítem particular, [es] el tema de cómo se cumple con los tres principios del retorno y la reubicación, que son: la voluntariedad, la seguridad y la dignidad. Entonces más que tener sistematizados como se verifican, se está replanteando cual es la forma adecuada en que efectivamente se implementa ese principio. Si yo redefino cual es la forma en que se materializo o vuelvo un poco más tangible la verificación o la materialización de ese principio, pues en ese orden de ideas ni siquiera es un tema de que yo ajuste solo la información, si yo ajusto el concepto, tengo que ajustar de ahí en adelante todo lo que este asociado a ese concepto. Entonces actualmente, por ejemplo, hablamos de las actas de voluntariedad, si tu te das cuenta la evaluación más o menos plantea que uno continua con esas actas de voluntariedad, entonces nosotros estamos analizando si realmente ese es el instrumento adecuado con el cual yo verifico el cumplimiento del principio o debo replantearme el instrumento con el cual lo verifico" (Asesora para la formulación de política pública, Subdirección General-UARIV).

> "… nosotros tenemos un sistema de información que se llama el MAARIV, es como un sistema para la dirección de reparación donde se desarrolla más técnicamente los pasos de un retorno (…) ahora estamos hablando de una semaforización del país, hay unas doce variables que vienen de un subcomité nacional de prevención y protección que hace parte del SNARIV, ellos toman una decisión, aprueban una propuesta que hacemos desde retornos y ya se incorpora en el sistema de información y ya hoy estamos hablando del semáforo y se ha venido socializando, no son todos, es un ejemplo, porque todo este cambio implica sistema de gestión, no solamente es incorporarlo sino explicárselo al resto de la entidad, que implica todos estos cambios, pero ese de semaforización ya está avanzando" (Funcionario Grupo de Retornos y Reubicaciones, UARIV).

La recomendación (o por lo menos su título) es demasiado abstracta, por lo que puede decirse que se están realizando acciones en esa dirección. Sin embargo, en cuanto a las acciones puntuales que se sugirieron como modo de implementación de la recomendación, únicamente se han implementado acciones concretas en el tema de alertas de seguridad (1 de 6).

Los demás puntos de la recomendación dependen de la operacionalización que se realice según la redefinición conceptual que trata la primera recomendación, por lo que su implementación a corto plazo, tal como se vio, es poco probable. Entonces, está recomendación no ha sido implementada, ni tampoco refleja efectividad del

proceso evaluativo, por lo que ninguna de las tres recomendaciones de esta evaluación fue realmente implementada en los términos propuestos.

6.5 Conclusiones

En general se observó un buen cumplimiento de los estándares profesionales de calidad dentro de esta evaluación. Sin embargo, existieron problemas en cinco estándares: participación de los actores relevantes, consideración de evaluación conjunta, enfoque asociativo, lógica de la intervención, y evaluación dentro del plazo y presupuesto asignado.

Con relación a los patrones de interacción y a la coordinación que se dio dentro del proceso de evaluación, se encontraron fallas en cuanto al hallazgo e involucramiento de todo el conjunto de actores pertinentes como usuarios previstos dentro del proceso de evaluación. Esto está en correspondencia con las fallas relacionadas con los estándares que tienen una naturaleza de reglas de frontera. A pesar de esto, los actores más centrales en el proceso de implementación de la política (UARIV y DPS) si fueron incluidos en el equipo de trabajo conjunto.

Se encontró que tanto el equipo evaluador como la DSEPP-DNP tuvieron una buena disposición de trabajo con el resto de los miembros del Comité de Seguimiento, esto se ve reflejado, entre otras cosas, en su elevado grado de reactividad frente a las demandas, preocupaciones y comentarios que los usuarios previstos realizaban a los productos de la consultoría, a la metodología seleccionada, y a la interpretación de los hallazgos y construcción de las recomendaciones.

En cuanto a efectividad, se encuentra que no se ha implementado la primera y más importante recomendación, y es poco probable que suceda del modo en que es propuesta por la evaluación. Si bien se está “trabajando” en la dirección de una redefinición conceptual del acompañamiento estatal en los procesos de retorno y reubicación, la evaluación es solo un insumo -entre muchos otros- dentro de este proceso, y con el tiempo ha ido perdiendo cada vez más relevancia. La segunda recomendación es una justificación de una decisión previa. Y la tercera, al depender de la redefinición conceptual que plantea la primera, no ha sido tampoco implementada.

La última versión del PTIR fue entregada el 25 de octubre de 2016, posterior a esta fecha se elaboró el proyecto de decreto modificatorio al Decreto 1084 de 2015, en donde se buscaría modificar la institucionalidad de la política que fue recopilada de la Ley 1448 de 2011 y el Decreto 4800 de 2011. Los funcionarios de la UARIV resaltan que están, a 2019, en ese proceso de redefinición de la normatividad y del diseño de la política, sin embargo, a parte de este proyecto de decreto, a julio de 2019 no existe indicio de otros cambios institucionales de la política, posteriores a

su evaluación que estuvieran a cargo de la misma UARIV. Lo más probable es que primero se modifique la Ley 1448 de 2011 en el Congreso de la República, y luego se entre en una etapa de modificación de las especificidades institucionales de la Política de Retornos y Reubicaciones. Por lo que a pesar de que hubo apertura a cambios instituciones en el nivel técnico, a nivel directivo y político esto aún no ha sucedido.

7. Efectividad de la Evaluación Institucional y de Resultados del Programa de Reparación Colectiva.

7.1 Presentación del Programa de Reparación Colectiva

El Programa de Reparación Colectiva (PRC) constituye una línea de acción particular dentro del componente de Reparación Integral, de la Política Pública para la Prevención, Protección, Atención, Asistencia y Reparación Integral a las Víctimas del Conflicto Armado. Este programa está dirigido a sujetos de reparación colectiva (SRC) que hayan sufrido un daño colectivo ocasionado por la violación de los derechos colectivos; la violación de los derechos individuales de los miembros del colectivo; o un impacto colectivo por violación a derechos individuales en el marco del conflicto armado (UARIV, 2015, p. 6).

La reparación colectiva es entendida como un derecho del que son titulares grupos, pueblos y organizaciones sociales que han sido afectadas, por lo que busca garantizar una reparación justa, adecuada y proporcional, de tal manera que las víctimas colectivas hagan un tránsito hacia la paz duradera (p. 7). Se refiere a “el conjunto de medidas de restitución, indemnización, rehabilitación, satisfacción y garantías de no repetición en los componentes político, material y simbólico tendientes a resarcir los daños sufridos por los SRC” (Centro Nacional de Consultoría –CNC-, 2016a, p. 1).

Al igual que los dos anteriores casos, la principal institucionalidad alrededor del PCR está incluida en la Ley 1448 de 2011 (art. 151) y en el Decreto 4800 de 2011. Los objetivos que contempla son:

- Reconocimiento y dignificación de los sujetos de reparación colectiva.
- Reconstrucción del proyecto de vida colectivo y/o planes de vida y/o proyectos de etnodesarrollo.
- Recuperación psicosocial de las poblaciones y grupos afectados.
- Recuperación de la institucionalidad propia del Estado Social de Derecho.

- Promoción de la reconciliación y la convivencia pacífica.

El PRC se implementa a través de Planes Integrales de Reparación Colectiva (PIRC), los cuales son formulados de manera participativa y con enfoque territorial dentro de los Comités Territoriales de Justicia Transicional. Estos PIRC pueden incluir diversos tipos de acciones dependiendo del diagnóstico que se realice del daño sufrido por parte del SRC, sin embargo, las principales acciones que contempla el programa (Mapa de política, p. 6) son: Restitución de territorios y retorno y/o reubicación para sujetos colectivos; Reconocimiento y dignificación de los sujetos de reparación colectiva mediante la reconstrucción y preservación de la memoria histórica; Articulación de medidas materiales de reparación colectiva; Cogestión, coordinación y priorización para el goce efectivo de derechos; Restitución de derechos de grupos sociales y políticos; Rehabilitación comunitaria y reconstrucción del tejido social; Reparación de las afectaciones ocasionadas al ambiente.

7.2 Caracterización de las situaciones de acción del proceso de evaluación del Programa de Reparación Colectiva

7.2.1 Reglas de alcance y resultados potenciales

Con respecto al cumplimiento de los estándares que se dirigen a establecer reglas de alcance dentro del proceso de evaluación, en primer lugar, frente a la fundamentación y finalidad de la evaluación, en los documentos de diseño de la evaluación se resalta la importancia coyuntural del objetivo de identificar: 1) el grado de articulación entre las entidades que están a cargo de la formulación e implementación del PRC a nivel nacional y territorial; 2) El grado de cumplimiento de los objetivos establecidos en el Decreto 4800 de 2011 (DSEPP-DNP, 2015e, p. 3), en el sentido en que la Reparación Colectiva es considerada como una herramienta importante para el posconflicto y para la reconstrucción de las comunidades a partir de la implementación de los acuerdos de paz.

La evaluación combinó en su alcance elementos de evaluación de resultados con evaluación de capacidad institucional, teniendo en cuenta que los distintos actores consideran el PRC como una política incipiente. Se esperó que los resultados de esta evaluación constituyeran un conjunto de oportunidades que fueran aprovechadas por parte de los tomadores de decisión (CNC, 2016b, p. 7), los cuales, para el caso del PRC, corresponden a los miembros del Subcomité de Reparación Colectiva (art. 241 del Decreto 4800 de 2011): Ministerio de Justicia y del Derecho, el Ministerio de Hacienda y Crédito Público, Departamento Administrativo para la Prosperidad Social, DNP, UARIV, Ministerio de Salud y Protección Social, Ministerio de Educación Nacional, Ministerio de Cultura, Ministerio de Agricultura y Desarrollo Rural, Ministerio de Vivienda, Ciudad y Territorio, Ministerio del Trabajo.

Puntualmente, el objeto de la consultoría fue:

> "Realizar una evaluación institucional y de resultados del Programa de Reparación Colectiva, que permita identificar las acciones a mejorar sobre la articulación y la capacidad institucional de las entidades a cargo de su formulación e implementación, a nivel nacional y territorial, para el cumplimiento de sus objetivos de acuerdo con lo establecido en el Decreto 4800 de 2011" (DSEPP-DNP, 2015e, p. 3).

Por un lado, la evaluación institucional se direccionó a la determinación de si la institucionalidad del PRC lleva al logro de los objetivos del programa, y del modo en que las normas y el contexto social y político inciden en su implementación (CNC, 2016b, p. 24). Para esto, se realizó un análisis de factores exógenos (contexto) y endógenos (capacidad institucional), de las fortalezas, obstáculos y debilidades, y de cómo estos se relacionan con los resultados. Por otra parte, el componente de evaluación de resultados se direccionó a identificar el avance en el cumplimiento de los objetivos del programa, particularmente abordando la implementación de la ruta y de las medidas de reparación colectiva, y la asociación entre la intervención y los resultados del programa (DSEPP-DNP, 2015e, p. 57).

Adicional a los componentes institucional y de resultados, se incluyó un tercer componente que brinda un marco para la construcción de las recomendaciones de la evaluación. En este, se estable que las recomendaciones deben dirigirse hacia la visión estratégica, la estructura institucional y organizacional para el mejoramiento de la implementación del PRC en los distintos niveles de gobierno, y para la obtención de mejores resultados (p. 58). También son especificados doce objetivos específicos dentro del pliego de condiciones definitivo (p. 55), los cuales guardan coherencia con los componentes descritos y con la finalidad y alcance de la evaluación. Con todo lo anterior, los estándares de fundamentación y finalidad de la evaluación, objeto y alcance de la evaluación, objetivos específicos y desarrollo de capacidad fueron adecuadamente direccionados y satisfechos dentro de este proceso de evaluación.

Al igual que en los anteriores procesos de evaluación, dentro de los documentos de diseño de la evaluación se observa el establecimiento de reglas y normas con respecto a la claridad, la pertinencia, la focalización y la orientación hacia la puesta en práctica de las conclusiones y recomendaciones de la evaluación. En esta evaluación también se partió de una primera propuesta de recomendaciones a la luz de los hallazgos de la evaluación, por parte del equipo evaluador, para posteriormente pasar por un doble proceso de mejora, primero a través de una mesa de expertos, y, posteriormente, con las consideraciones de los miembros del Comité de Seguimiento.

En la mesa de expertos se validaron las recomendaciones a partir de los criterios de (CNC, 2016c, p. 6): pertinencia; claridad; viabilidad institucional, técnica y

presupuesta; responsable claro y ajustado a sus competencias; tiempo de implementación; y nivel de criticidad. Las recomendaciones fueron presentadas en el documento PTIR, en donde se refleja una clara conexión entre estas y los hallazgos de la evaluación, y son formuladas adecuadamente. En este sentido, los componentes del estándar de oportunidad, pertinencia y utilización que son reglas de alcance fueron satisfechos completamente.

7.2.2 Reglas de frontera y participantes

Únicamente se encontró evidencia de la participación de la UARIV, el DPS y el DNP en el Comité Técnico de Seguimiento[61], esto significa que de once entidades que conformar el Subcomité de Reparación Colectiva (instancia de coordinación interinstitucional a través de la que se formula y se implementa el PRC), únicamente tres entidades tuvieron participación en el diseño y en la ejecución de la evaluación de este programa. Así mismo, se observa que, a lo largo de la presentación de las recomendaciones en el PTIR, se establece como responsables de su implementación a estas entidades del Subcomité de Reparación Colectiva a pesar de que no fueron incluidas en el proceso de evaluación. Esto quiere decir que los estándares de participación de los actores relevantes, enfoque asociativo y consideración de evaluación conjunta no fueron satisfechos en este proceso de evaluación.

Por otra parte, en cuanto al equipo de la evaluación, se seleccionó al Centro Nacional de Consultoría (Resolución 3903 de 2015) para la realización de la evaluación. La modalidad de selección fue a través de concurso de méritos abierto (DSEPP-DNP, 2015e, p. 4), en donde los proponentes debieron pasar por un proceso de habilitación de sus propuestas, para luego proceder a una etapa de calificación de las propuestas metodológicas. Dentro de los requisitos técnicos habilitantes, se incluyó la postulación de un equipo mínimo conformado por (p. 17): un Director del proyecto con nivel educativo mínimo de maestría y con tres años de experiencia en dirección o coordinación de evaluaciones o estudios relacionadas con programas de ciencias sociales; un Experto en Población desplazada con mínimo maestría y dos años de experiencia relacionada; y un analista cualitativo con mínimo especialización en ciencias sociales y dos años de experiencia relacionada.

Dentro de los criterios de calificación de las propuestas estuvieron: el enfoque metodológico, la experiencia del equipo de trabajo, y la formación académica y las publicaciones técnicas y científicas del equipo de trabajo (p. 25). Todo el proceso de contratación, comenzando por el aviso de convocatoria, la publicación de los informes de habilitación y de calificación, hasta la publicación del contrato, se realizó en el portal SECOP. Todo lo anterior satisfizo el estándar de equipo de evaluación.

[61] Esto con base en la respuesta a derecho de petición del DNP N° 20193600428271.

Frente al estándar de recursos se observa que, por un lado, el DNP dispuso un presupuesto para la evaluación de $496.339.640, los cuales pagó al consultor del siguiente modo: un 20%, una vez se haya entregado a satisfacción el informe metodológico; un 20%, con entrega a satisfacción del primer informe de avance de trabajo de campo; un 20%, con entrega a satisfacción del informe final de trabajo de campo; un 25%, con entrega a satisfacción del informe final de la evaluación; y un 15%, con entrega a satisfacción del PTIR (DNP, 2015b, p. 7). Por parte del equipo evaluador, este se comprometió a disponer del equipo de trabajo ofertado en la propuesta, de acuerdo con los requerimientos técnicos (DSEPP-DNP, 2015e, p. 3), todo esto satisfizo el estándar concerniente.

7.2.3 Reglas de elección y acciones

Con respecto a las acciones realizadas por los participantes en la situación de acción de diseño de la evaluación, al observar el grado de cumplimiento del estándar de evaluabilidad se encuentra que, en esta etapa de diseño no se contó con una definición clara de la lógica de la intervención a través de metodologías como marco lógico o cadena de valor (CNC, 2016a, p. 9; CNC, 2016c, p. 12). Adicionalmente, el programa no contó con resultados operacionalizados en un conjunto claro de indicadores y metas en ese momento de la evaluación.

Sin embargo, se realizaron consideraciones alrededor de la viabilidad de la evaluación -particularmente al definir la metodología de la evaluación a partir de características singulares del PRC- tales como: la existencia de múltiples actores en diferentes niveles de gobierno; la interacción del PRC en todas sus etapas con beneficiarios diversos y complejos; la necesidad de una aproximación flexible dada la diversidad del PRC; la necesidad de aprovechar el conocimiento de actores situados en las prácticas institucionales del programa (CNC, 2016b, p. 25). Con todo esto, se estableció que lo más viable era la realización de una evaluación completamente cualitativa.

También se realizó un análisis de los riesgos previsibles que pudieron afectar el "equilibrio contractual" dentro de los documentos de diseño de la evaluación. A cada riesgo se le realizó una descripción, su fuente u origen, la etapa en la que surge, la consecuencia de su ocurrencia, su probabilidad, su impacto, su prioridad, su tratamiento o control y los encargados de implementar ese tratamiento o control (DSEPP-DNP, 2015e, pp. 48-50). Entre estos riesgos se encuentran: la escasez de oferta de firmas consultoras que cumplan con los requisitos; los retrasos en la constitución y presentación de las garantías que amparar el cumplimiento del contrato por parte del consultor; los retrasos o incumplimientos en la entrega de productos por parte del consultor; las demoras por parte del DNP en la aprobación de los productos desarrollados por el consultor; las dificultades o impedimentos para acceder a la zona donde se realiza el trabajo de campo por parte del contratista por razones de orden público. Con todo esto, solamente puede considerarse como parcialmente satisfecho el estándar de evaluabilidad.

Teniendo presente que se considera al PRC como una política pública en germinación (CNC, 2016b, p. 8), la única normatividad referente a la intervención que es analizada dentro de la evaluación es la Ley 1448 de 2011, el Decreto 4800 de 2011, el Decreto 4633 de 2011, el Decreto 4634 de 2011 y el Decreto 4635 de 2011 (Estos últimos tres corresponden a la reglamentación de la Ley 1448 de 2011 para los grupos étnicos), y el Decreto 1084 de 2015. Así mismo, al ser los Planes Integrales de Reparación Colectiva (PIRC) el medio por el cual se implementa el PRC, al ser estos construidos a nivel local y con la participación de los SRC, y al ser tomados como casos de análisis de la evaluación a 24 SRC en 9 departamentos (CNC, 2016b, p. 91), se puede manifestar que el estándar de coordinación y alineamiento también fue satisfecho.

Frente a la selección del enfoque metodológico de la evaluación -dentro de los lineamientos del contrato y del pliego de condiciones definitivo- se observa coherencia entre la finalidad de la evaluación, su alcance, las preguntas orientadoras de la evaluación, y el marco general que se le da al consultor para la construcción de su propuesta metodológica, esto significó que si bien el DNP conservó la última palabra en la aprobación de todos los componentes de la metodología, la discrecionalidad del consultor para su formulación fue considerable.

Particularmente, dentro del pliego de condiciones definitivo, se solicitó especificar una metodología de análisis cualitativo, incluyendo las técnicas a utilizar, los supuestos subyacentes, y la pertinencia y justificación de estas para: el mapa de actores y relaciones interinstitucionales del PRC; un análisis de la articulación institucional de las entidades del SNARIV a nivel nacional y territorial identificando factores de éxito y puntos críticos; un análisis de la capacidad institucional de las entidades; la identificación de variables de análisis para responder a las preguntas orientadoras; la identificación de fuentes de información primaria y secundaria; la contribución de la implementación de la ruta de reparación colectiva sobre sus objetivos (DSEPP-DNP, 2015e, p. 27).

Así mismo, se les requirió a los proponentes una estrategia metodológica para la integración de los hallazgos de los componentes institucional y de resultados de la evaluación, el cual debió orientarse a la identificación de las relaciones entre la institucionalidad y los resultados que se encontraran (p. 27). Adicional a esto, se establecieron lineamientos acerca de las unidades de análisis a tener en cuenta, los aspectos a evaluar, justificación técnica de la metodología, levantamiento de información de campo (pp. 58-63). En consonancia, el equipo de evaluación seleccionó el Sistema de Análisis y Desarrollo de la Capacidad Institucional -SADCI- (CNC, 2016b, p. 10) como metodología de evaluación institucional.

Por su parte, para la evaluación de resultados se observa una metodología de caracterización cualitativa enfocada hacia las narrativas, las razones y los puntos de vista de los actores del PRC (p. 91). El único componente del estándar que presenta falencias es que al no poseer una lógica de intervención tampoco se podía hacer distinción entre los distintos niveles de resultados. Sin embargo, esto puede

llegar a ser justificable gracias a la naturaleza del programa y a la flexibilidad de los PIRC, por lo que el estándar de selección del enfoque y método fue satisfecho.

Se refleja también la utilización de criterios de evaluación. Particularmente se observa: coherencia, definición de un área de mejora y propósito, y viabilidad, como criterios de evaluación de la pertinencia de los objetivos del PRC (CNC, 2016c, p. 10). Se indaga (DSEPP-DNP, 2015e, p. 56) sobre la pertinencia entre competencias de entidades en relación con los diagnósticos realizados de los daños sufridos por los SRC; la sostenibilidad de la participación y contribución de la entidades del estado; la pertinencia del PRC con respecto a la capacidad institucional y fiscal del Estado; la pertinencia de las medidas de los PIRC con respecto a la reparación del daño de los SRC; el modo en que se puede hacer sostenible la financiación de los PIRC, entre otras. se realizaron 11 preguntas para el componente institucional, 19 preguntas para el componente de resultados, y 7 preguntas para el componente de recomendaciones de la evaluación. La coherencia entre preguntas y objetivos de la evaluación permite establecer que los estándares de cuestiones a responder y de selección de criterios de evaluación fueron completamente satisfechos.

Con relación a las acciones de los participantes en la situación de acción de ejecución de la evaluación, comenzando por la respuesta a las preguntas evaluativas a lo largo del informe final de la evaluación y del PTIR, la mayoría de las veces no se hace referencia a las preguntas cuando se presentan los hallazgos correspondientes, lo cual resta en claridad del análisis, sin embargo, únicamente faltó responder cuatro preguntas de las 37 formuladas en el pliego de condiciones definitivo, estas fueron: ¿Cómo se podría identificar que un sujeto colectivo está plenamente reparado? ¿Qué tipo de medidas podrían ser identificadas por los sujetos de reparación colectiva como más importantes o con mayores efectos en su proceso de reparación? ¿Qué implicaciones tiene para la comunidad medidas aprobadas que no son implementadas completamente, o implementadas extemporáneamente? ¿En qué medida la orientación del Programa relaciona a la población víctima y no víctima? Lo más probable es que estas no pudieran ser respondidas debido a la heterogeneidad de medidas de los distintos PIRC, y de los contextos en los cuales se implementan.

En el PTIR se encuentra mejor evidencia que en el informe final acerca de la separación y derivación lógica entre los hallazgos, las conclusiones y las recomendaciones. Al interior del primero se presenta cada recomendación como un particular subtítulo, en los cuales se realiza una descripción introductoria de los hallazgos y conclusiones de la evaluación que justifican y sustentan empíricamente la recomendación. Lo anterior satisfizo los estándares de respuesta a preguntas evaluativas y claridad del análisis, a pesar de que no sean estándares que se vean reflejados en el informe final de la evaluación, y que por lo tanto no tuvieron difusión externa.

El informe final no cuenta con un resumen ejecutivo. Al igual que en anteriores evaluaciones, dentro de los últimos productos requeridos al equipo evaluador, se encuentran el resumen ejecutivo y la ficha técnica de la evaluación como documentos separados del informe final (DSEPP-DNP, 2015e, pp. 68-69). Dentro del resumen (CNC, 2016d) se realiza una breve presentación de la evaluación, una presentación del PRC, y una descripción general de los hallazgos de la evaluación institucional, y de la evaluación de resultados. La ficha técnica (2016a), por su parte, incluyó una descripción del PRC, el objetivo de la evaluación, la metodología de la evaluación, los principales hallazgos y conclusiones, y la presentación de las recomendaciones, por lo que esta última tuvo mejor capacidad como herramienta de difusión. Con estos dos documentos se satisfizo el estándar de claridad y representatividad del resumen.

En lo concerniente a la descripción y análisis del contexto de la intervención, se observa que, al seleccionar una metodología únicamente cualitativa, la evaluación realizó un análisis profundo del contexto en sus distintos componentes. Entre las variables de contexto que son analizadas están: el sentido de las acciones, prácticas y códigos compartidos de los actores y las instituciones (2016b, p. 24); la participación de los actores implicados y su coordinación (2016a, P. 3); la institucionalidad concerniente a competencias y responsabilidades de los actores (p. 3); el grado de organización y empoderamiento de las comunidades (p. 5); las condiciones contextuales de los territorios en los cuales se asientan las comunidades (p. 5); la situación actual de la implementación de los acuerdos de La Habana, su relación con la reparación colectiva y el potencial que esto tiene para la reparación de los daños ocasionados en el marco del conflicto armado (2016b, P. 7); el contexto económico; el contexto cultural (DSEPP-DNP, 2015e, p. 60). Lo anterior satisfizo el estándar de contexto de la intervención.

En relación con la lógica de la intervención que debe ser incluida en los productos de la consultoría, tal como se mencionó dentro de las acciones correspondientes a la etapa de diseño de la evaluación, una vez esta fue ejecutada uno de sus hallazgos fue la necesidad de realizar una mejor delimitación del programa a nivel de medio, componente o estrategias a partir de las cuales alcanzar los objetivos, a partir de una herramienta sintética como la de cadena de valor (CNC, 2016c, p. 12). La ausencia de lo anterior está relacionada con el hecho de que existiera una concomitancia entre los procesos de formulación (y su necesaria coordinación interinstitucional) de herramientas metodológicas del PRC, la socialización del programa, y la implementación de este (CNC, 2016b, p. 118; 2016a, p. 3).

A pesar de la limitación que implica la indeterminación de la relación que debe existir entre medios y fines, la evaluación realizó una descripción del PRC, enfocándose en sus objetivos, sus componentes, el marco institucional que lo sustenta (CNC, 2016b, p. 11), las medidas de reparación colectiva (p. 16), la ruta de reparación colectiva –incluyendo una descripción de cada una de sus fases- (p. 21). Por esta razón, si puede decirse que hubo descripción y valoración de la lógica de la intervención, por lo que el estándar correspondiente fue cumplido.

También se observa una adecuada explicación de la metodología de evaluación dentro del informe final de la consultoría. En esta explicación se profundiza acerca de la metodología para el componente institucional y el de resultados. Con respecto al primero, tal como se mencionó en el estándar de selección del enfoque y método, se utilizó la metodología SADCI (Sistema de Análisis y Desarrollo de la Capacidad Institucional, la cual se enfoca en analizar las fortalezas, los déficits de capacidad institucional y los principales cuellos de botella de las entidades participantes en la intervención evaluada (CNC, 2016b, p. 21; 2016a, p. 1). Esta metodología deductiva se combinó conceptualmente con categorías que fueron surgiendo de modo inductivo a medida que se desenvolvía la recolección de datos cualitativos, para de este modo generar líneas de análisis sobre los datos, los cuales se basaron en relatos, códigos y significados de las personas seleccionadas como muestreo intencional (2016b, p. 25).

La metodología del componente de evaluación de resultados se enfocó en indagar sobre los avances del programa, analizando narrativas, razones y puntos de vista que tienen los distintos actores, y particularmente los SRC, acerca de los resultados que ha tenido el programa (p. 91). También se enfocó en comprender el modo en el que se ha desenvuelto el programa y la manera en la que se ha ido obteniendo los objetivos del mismo. Para esto, se utilizó la técnica de grupos focales (p. 93), generando discusiones alrededor de: el avance en el cumplimiento de las metas del PRC; la forma en la que se han desarrollado cada una de las fases de la ruta de reparación colectiva; los efectos reparadores relacionados con la implementación de la ruta; y la relación entre la oferta institucional y el desarrollo territorial (p. 93).

En este proceso de evaluación se realizaron 109 entrevistas semiestructuradas en 26 municipios ubicados en 10 departamentos para el componente de evaluación institucional (2016a, p. 2). Estas entrevistas se presentan de manera desagregada en tablas que indican las entidades a las que se les realizaron, y el nivel de gobierno al que corresponden (2016b, p. 27). Por su parte, los datos del componente de resultados consistieron en grupos focales a 24 SRC en 9 departamentos, que incluyeron 4 casos nacionales realizados en Bogotá (2016a, p. 2). Estos 24 sujetos fueron caracterizados por departamento, municipio, nombre, tipo de sujeto, su fase en la ruta de reparación colectiva y el número de asistentes al grupo focal (2016b, p. 91). Esto se complementó con estudios de caso que buscaron ilustrar y ejemplificar algunos de los hallazgos más relevantes de la evaluación (p. 29).

La presentación de la metodología, por último, también incluyó menciones acerca de sus límites. En primer lugar, se señala que los principales efectos reparadores del programa son el reconocimiento como víctimas de los SRC, la generación de confianza en el estado, el empoderamiento de los SRC, y la contribución a los objetivos de recuperación psicosocial, sin embargo, debido a su grado de abstracción, su medición (y seguimiento) resulta difícil (2016a, p. 4). Adicionalmente, se señalan retos relacionados con la traducción de las narrativas y opiniones de los SRC a referentes que permitan construir gradaciones o parámetros

de medición de logros y avances. Y, por último, el reto de establecer una adecuada consideración del significado de la reparación colectiva en un contexto de justicia transicional (2016b, p. 11). Toda esta descripción cumplió con los componentes de los estándares de explicación de la metodología, validez y fiabilidad de las fuentes de información, y de mención de cambios y límites de la metodología de evaluación.

Dentro del pliego de condiciones y en el contrato de consultoría se establecieron lineamientos claros orientados hacia la consulta y la protección de los implicados. Dentro de estos lineamientos está la obligación del DNP de anonimizar la información producto de la consultoría, que tenga relación con los derechos a la intimidad, habeas data, buen nombre, y que en general tenga protección legal (DSEPP-DNP, 2015e, p. 5; DNP, 2015b, p. 6). Lo que satisfizo el estándar de consulta y protección de implicados.

Tampoco se realizan menciones acerca de las discrepancias entre los miembros del equipo evaluador, sin embargo, se describe en detalle el modo en que estas discrepancias fueron solucionadas, al respecto se relata la realización de un taller presencial de todos los miembros del equipo evaluador, en donde cada uno presentó recomendaciones con base en los hallazgos, y los demás realizaron una labor crítica, de validación o rechazo (CNC, 2016c, p. 5), esto refleja el cumplimiento del estándar de mención de discrepancias entre miembros del equipo evaluador.

Finalmente, para cerrar la medición del cumplimiento de los estándares que corresponden a reglas de elección, en los términos del contrato, el informe final de evaluación y el PTIR tuvieron fecha de entrega del 2 de junio y del 5 de julio de 2016 respectivamente. A pesar de que no se encontró algún documento de adición al contrato (como un otrosí) en el portal SECOP, la fecha del informe final es del 23 de septiembre de 2016, y del PTIR únicamente se indica a octubre de 2016. Esto significa un retraso considerable frente a lo que estaba previsto en el contrato, por lo que el estándar de realización de la evaluación en el plazo y el presupuesto asignados no fue satisfecho.

7.2.4 Reglas de agregación y control

Las reglas de agregación y el control ejercido por los participantes frente a la conexión entre acciones y resultados potenciales, es bastante similar a las de las evaluaciones anteriores. Se establecen reglas de control de calidad como la entrega de informes técnicos por parte del equipo evaluador cada vez que le sean solicitados en un plazo no mayor a 8 días hábiles (DSEPP-DNP, 2015e, p. 4); la garantía del cumplimiento de las obligaciones a favor del DNP a partir de mecanismos de cobertura de riesgo como póliza, patrimonio autónomo o garantía bancaria (p. 51); clausula penal pecuniaria por valor al 30% del contrato como garantía del cumplimiento de las obligaciones surgidas del contrato, de la calidad del servicio y del pago de salarios, prestaciones sociales e indemnizaciones laborales; y multas en caso de mora o incumplimiento parcial de cualquiera de las obligaciones del consultor (DNP, 2015b, Cláusula 19).

Del mismo modo, la supervisión del contrato estuvo a cargo del DNP a partir del establecimiento de una secretaría técnica dentro del Comité de Seguimiento, la cual debió: aprobar todos los instrumentos de recolección de datos, y todos los ajustes metodológicos y técnicos (p. 66); poner a consideración del Comité de Seguimiento los productos del consultor, recibir los comentarios que este haga a los productos, consolidarlos, depurarlos y darles el tratamiento respectivo (p. 53); y una vez los comentarios fuesen integrados por parte del consultor, pudo entregar el recibo a satisfacción de los productos para proceder al pago correspondiente (DNP, 2015b, p. 4). Esto significó un cumplimiento de todos los componentes de los estándares de control de calidad y de inclusión de observaciones de los implicados.

7.2.5 Reglas de posición

Los comentarios a los productos realizados por el Comité de Seguimiento tuvieron previsto ser discutidos (DSEPP-DNP, 2015e, p. 53), por lo que el equipo evaluador tuvo cierta discrecionalidad de defender sus productos en dado caso que no estuviera de acuerdo con los comentarios del Comité de Seguimiento. Sin embargo, y al igual que en las anteriores evaluaciones, los altos costos asignados a las posibles acciones que pudo realizar el DNP en dado caso que quisiera activar algunas de las cláusulas del contrato correspondientes a reglas de agregación, así como el hecho de que solo se cuente con una semana desde que se reciben los comentarios para la entrega de los productos corregidos (DNP, 2015b, p. 4), sirven a modo de incentivo para que el equipo evaluador prefiera sacrificar su independencia como evaluador antes de entrar en una discusión que fácilmente puede llevar a una interpretación de incumplimiento de los términos del contrato. Al observar el tipo de comentarios que fueron realizados por el Comité de Seguimiento al informe final y al PTIR, se tiene lo siguiente:

> "…no se observó una conexión entre los sujetos de reparación colectiva y las medidas de reparación, tal que no se identifican detalles como: cuál fue su grado de aceptación, en qué fase de la ruta se encuentran, en general qué solicitan para su reparación, qué consideran fundamental para sentirse reparados, qué consideraciones tienen del programa…" (DSEPP-DNP, 2016b, p. 3)

> "…no hace uso adecuado de las categorías de análisis para determinar el grado de articulación, coordinación institucional del programa (…) no se presentan análisis sobre los resultados en términos de instrumentos, decisiones, lineamientos de política, acuerdos interinstitucionales y demás desarrollo de política pública…" (p. 3)

> "…desarrollar a mayor profundidad las recomendaciones. A juicio del comité ninguna aclara los factores que deben ser tenidos en cuenta para su

> implementación, y la mayoría no precisan qué se está recomendando, ello por la falta de desarrollo o el nivel demasiado general" (2016c, p. 4)
>
> "…las recomendaciones están muy centradas en la evaluación institucional, se deben incluir y mejorar algunas que se refieran a la evaluación de resultados; al igual que algunas son repetitivas y no es claro cómo se debe realizar su implementación" (p. 6)

En general, los comentarios acerca de los hallazgos y conclusiones apuntan a la inclusión de diversos temas que fueron dejados de lado, mientras que los comentarios acerca de las recomendaciones piden profundizar en los mecanismos de implementación. No se encontró ningún comentario dirigido a la modificación de la esencia de las valoraciones o las recomendaciones de la evaluación. Además, el mismo equipo de evaluación reconoce explícitamente que la responsabilidad final de sus conclusiones y recomendaciones recayó exclusivamente en ellos (CNC, 2016c, p. 5). Este reconocimiento explícito, permite establecer que en esta evaluación se cumplieron los estándares de independencia de los evaluadores y de proceso evaluativo libre y abierto.

7.2.6 Reglas de información

Toda la documentación acerca de la finalidad y la expectativa con respecto a la consultoría de este proceso de evaluación se pudo encontrar en el portal del SECOP. Al igual que en evaluaciones anteriores, se pueden encontrar los documentos de estudios previos, proyecto de pliego de condiciones y pliego de condiciones definitivo, con los cuales se cumplió con todos los componentes del estándar de documentación acerca de la finalidad y la expectativa de la evaluación.

Los productos principales de la evaluación, el informe final y el PTIR, contaron con adecuados lineamientos que garantizaron que su contenido fuera comprensible para los usuarios previstos, y para los demás actores involucrados en el PRC (DSEPP-DNP, 2015e, p. 67). En cuanto a su difusión, estos productos fueron socializados internamente ante el Comité de Seguimiento, y ante los directivos de las entidades que lo conformaron (p. 4), con el objeto de realizar las retroalimentaciones que se consideraran necesarias.

Estos productos fueron acompañados de herramientas diseñadas para la difusión de la evaluación, tales como la ficha técnica de la evaluación, una presentación extendida, una presentación simple, una infografía y un video con el resumen de los principales resultados de la evaluación (p. 69). Para efectos de difusión externa, al igual que el resto de las evaluaciones de la DSEPP, se publicó el informe final, el resumen ejecutivo y la ficha técnica en el portal de SINERGIA. De este modo, los estándares de informe de la evaluación y difusión de la evaluación fueron satisfechos.

7.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación del Programa de Reparación Colectiva

Tal como se observó en el estándar de participación de los actores relevantes, en este proceso de evaluación, la mayoría de los miembros del Subcomité de Reparación Colectiva no participaron como usuarios previstos dentro de un equipo de trabajo conjunto, a pesar de ser los encargados de implementar y formular el PRC. En este sentido, si bien se trabajó con la UARIV y el DPS como principales usuarios previstos, los cuales son los encargados de coordinar y liderar el PRC, su poder para implementar de las recomendaciones es, hasta cierto punto, limitado. Este problema es parcialmente reconocido dentro de la evaluación:

> "… otro inconveniente que sobresale en el análisis (y en las entrevistas) está relacionado con el lugar que ocupa la UARIV en el Sistema. Al ser una entidad del tercer nivel en la jerarquía del Estado, adscrita al sector de la Inclusión Social y la Reconciliación, la tarea de coordinar entidades de segundo y primer nivel, así como de otras ramas del poder público, resulta complicada. En el nivel nacional, la UARIV es la entidad encargada de coordinar operativamente el Sistema, cosa que no siempre resulta fácil dados los rasgos señalados. Desde el punto de vista de los funcionarios entrevistados, expresiones como el "no tener dientes", se usan precisamente para señalar las complicaciones que se generan al darle la labor de coordinación a una entidad con un nivel jerárquico menor al de las que tiene que convocar" (CNC, 2016b, p. 40).

A pesar de esto, la UARIV al tener el rol de secretaría técnica dentro del Subcomité de Reparación Colectiva, tiene la posibilidad de impulsar cambios en el programa. Adicionalmente, se pudo constatar que únicamente a una de las recomendaciones se le asignó como responsables de su implementación, actores que fueron totalmente ajenos a los miembros del Comité de Diseño y de Seguimiento de esta evaluación[62], y que todas las recomendaciones que fueron implementadas, sus responsables fueron incluidos como usuarios previsto. En el resto de los casos, el alcance de las recomendaciones está dentro de la esfera de influencia de la UARIV o el DNP, por lo que hubo un buen nivel de involucramiento de los usuarios que se consideraron necesarios.

Con respecto a la dinámica del equipo de trabajo conjunto, los funcionarios de la UARIV manifestaron una percepción positiva con respecto a la fluidez del trabajo entre las tres partes, y subrayaron el papel del DNP como canal de comunicación

[62] Correspondió a la recomendación N° 8 de ajustar la composición del Subcomité de Reparación Colectiva, la cual tuvo como responsable el Comité Ejecutivo del SNARIV.

con el equipo de evaluación, que “genera acciones de tal forma que evite la subjetividad a la hora de tener un relacionamiento entre el evaluador y el evaluado” (Asesora para la formulación de política pública, Subdirección General-UARIV). En las entrevistas se destaca la importancia de la supervisión técnica que realizan:

> “… digamos que somos fuente de información en un momento determinado al ser evaluados, pero el evaluador es independiente, es autónomo, aunque si tiene una revisión técnica por parte del DNP, por la experiencia que ellos tienen en términos de las evaluaciones, por ejemplo, cuando se va a seleccionar la muestra, ellos si hacen el seguimiento particular al método, eso lo hace el DNP como el evaluador indirecto, por decirlo así, el directo seria econometría” (Asesora para la formulación de política pública, Subdirección General-UARIV).

También se encontró evidencia acerca de la reactividad del evaluador y del DNP con respecto a las preocupaciones e inquietudes de los usuarios previstos (a pesar de que solo estuviera la UARIV), por lo que no se permitió que el interés del evaluador o del DNP dominara el proceso de ajuste de prioridades. Dentro de los temas en los que se fue particularmente reactivo se señaló la determinación de la necesidad de información en la etapa de diseño de la evaluación, la definición de universos y de la muestra, y la socialización de los resultados de la evaluación (Asesora para la formulación de política pública, Subdirección General-UARIV).

Acerca de las discusiones metodológicas del equipo de trabajo conjunto, estas se encontraron enmarcadas institucionalmente dentro de la aprobación o rechazo que el supervisor del contrato (DSEPP) debió realizar a todos los instrumentos de recolección de información, y a todos los ajustes metodológicos y técnicos que se presenten en el desarrollo de la consultoría (DSEPP-DNP, 2015e, p. 66). Así mismo, debido a la naturaleza netamente cualitativa de la metodología, esta sufrió transformaciones tanto por la supervisión del contrato, como por el mismo desarrollo del método, en donde los usuarios previstos eran fuente de información primaria:

> “…el uso de métodos cualitativos implica partir de un principio (Jack Goody, 1990, lo llama aproximación reflexiva a los datos y modelos) que reconoce de entrada una relación de ida y vuelta, entre la teoría y la práctica, o entre los supuestos iniciales y el mundo empírico (…) Este procedimiento fue precisamente el que se siguió para recoger, sistematizar y codificar la información. En primer lugar, como corresponde a los métodos cualitativos, hubo una relación de ida y vuelta, entre las categorías iniciales de la evaluación (i.e. arquitectura, coordinación institucional, recursos) y las categorías que surgen de los datos obtenidos en el terreno, o categorías emergentes. Por medio de talleres internos, el equipo diseñó un “árbol” de códigos estructurado a partir de las categorías iniciales de análisis, complementado con las categorías emergentes” (CNC, 2016b, p. 26).

> “En términos de la operacionalización de las categorías analíticas utilizadas, y dado su carácter eminentemente cualitativo, la evaluación apostó por un

> diseño que, si bien partía de una estructura predefinida, se adaptaba a las categorías que fueron surgiendo tanto de la revisión de la documentación consultada, como de las entrevistas con los funcionarios de las entidades y los miembros de los SRC" (CNC, 2016a, p. 1)

En este sentido, el mismo método cualitativo requirió de una reactividad por parte del equipo evaluador con respecto a la adopción del método en respuesta a inquietudes o preocupaciones de los usuarios en este tema en particular. Este enfoque reflexivo también fue utilizado para la elaboración de las recomendaciones de la evaluación y para el establecimiento de los hallazgos y las conclusiones (CNC, 2016c, p. 5), por lo que puede manifestarse que los usuarios previstos también fueron involucrados en estos escenarios:

> "…las recomendaciones que propuso el equipo consultor contienen implícito el diálogo con las recomendaciones de las personas involucradas directamente con el programa. Obviamente, la responsabilidad final es del equipo consultor. Este diálogo enriqueció el ejercicio, más no lo determinó" (p. 5).

Por último, no se encontró ninguna evidencia de la planificación o realización de seguimiento a la implementación de los hallazgos de la evaluación, o de la búsqueda proactiva del uso por parte del equipo evaluador, más allá de los espacios de socialización que se le requirieron.

7.4 Análisis de la efectividad de la evaluación del Programa de Reparación Colectiva

La evaluación produjo un total de once recomendaciones. La primera fue revisar y redefinir los componentes del PRC, a partir de desagregar los objetivos que existen en la normatividad vigente, en componentes y estrategias, y que estas sean a su vez desagregadas en acciones que hagan las veces de medios para alcanzar los objetivos del programa (CNC, 2016c, p. 13). Para esto, se sugieren criterios de concreción en el alcance de los componentes; concreción en los conjuntos de acciones de cada componente; asociación de acciones con mecanismos concretos de implementaciones; asociación de acciones con oferta institucional; especificación de responsables de implementación de acciones; agregar indicadores de resultados a acciones (p. 13).

Esta recomendación tiene un grado considerable de abstracción, en el sentido en que muchas acciones pueden ser consideradas como una reformulación. En términos generales, tanto en la UARIV como en la Procuraduría General de la Nación (la cual hace parte de la Comisión de Seguimiento y Monitoreo al cumplimiento de la Ley 1448 de 2011, ver art. 201) manifestaron que esta recomendación ya se había realizado. Adicionalmente, tanto en el Proyecto de

Decreto Modificatorio del Decreto 1084 de 2015, con el que contaba la UARIV en julio de 2018 (DPS, 2018), como en la Resolución 3143 de 2018 de la UARIV, se incluyen modificaciones a los componentes y acciones del PRC. En cuanto al proyecto de decreto se establece:

> "… a través de las siguientes acciones específicas con los sujetos de reparación colectiva: 1. Reconstruir los proyectos colectivos, planes de vida o proyectos de etnodesarrollo (…) 2. Recuperar y restablecer el tejido social y las relaciones de confianza (…) 3. Dignificar y recuperar la memoria histórica (…) 4. Fortalecer las capacidades políticas, ciudadanas y de incidencia institucional territorial y nacional (…) 5. Contribuir al restablecimiento de confianza entre los ciudadanos y las instituciones del Estado que pueda facilitar la implementación del programa de reparación colectiva (…) 6. Promover la reconciliación, los pactos de convivencia al interior y con el entorno" (DPS, 2018, p. 13).

Por su parte, en la Resolución 3143 de 2018 se aclara:

> "ARTÍCULO SEXTO. PROCESOS DEL PROGRAMA ADMINISTRATIVO DE REPARACIÓN COLECTIVA. El Programa Administrativo de Reparación Colectiva estará constituido por tres procesos paralelos: 1) la ruta de reparación colectiva, que contribuye a la reparación de los daños colectivos, 2) el fortalecimiento del sujeto colectivo, que contribuye a instalar capacidades de autogestión para su desarrollo; y 3) el proceso de gestión de oferta, comprendido como las acciones adicionales que, sin ser competencia directa de la Unidad para la Atención y Reparación Integral a las Víctimas, establecen condiciones mínimas para el desarrollo del proceso de reparación colectiva".

Si bien es en el proyecto de decreto que están las modificaciones al PRC que más se aproximan a lo sugerido por la recomendación, y que únicamente se ha implementado lo que está en la resolución, puede decirse que la UARIV ha realizado esfuerzos significativos en el sentido de la recomendación. Por esta razón, se puede decir que el grado de implementación de la primera recomendación refleja el uso de la evaluación, así como la efectividad de esta.

La segunda recomendación fue desarrollar la dimensión colectiva de cada una de las medidas de reparación, lo que requirió de un esfuerzo conceptual que tuviera en cuenta que cada una de las medidas de reparación en su dimensión colectiva debía tener un propósito, un desarrollo basado en tipos de medidas agrupados por acciones o partes de un programa, y tener claros sus mecanismos de implementación (CNC, 2016c, p. 13). El grado de implementación de esta recomendación se respondió de manera conjunta con la tercera recomendación por parte de los entrevistados, a razón de su similitud. Esta consistió en clarificar los límites y las relaciones complementarias de la reparación colectiva con las medidas de desarrollo, esto, a partir de avanzar en la teorización del nexo causal que existe entre el daño sufrido por la comunidad, y las medidas de reparación que existan en

los PIRC, lo cual se sugería a partir de una tipología de los daños y las medidas que pueden diseñarse para hacerles frente (p. 16). En cuanto a su implementación, se manifestó:

> "… digamos estas dos incluso están relacionadas porque lo que se redefinió en el programa es el tema de que las acciones que quedaban definidas en los PIRC efectivamente dieran una respuesta a los daños colectivos sufridos por el sujeto. En ese orden de ideas, en la medida en que las acciones van encaminadas a resarcir los daños, pues como se desligan un poco de esas deudas históricas de desarrollo, que eran de alguna manera las que estaban quedando en los primeros planes de reparación colectiva que se formularon. No solo quedaban acciones relacionadas con lo que el sujeto colectivo había sufrido por el conflicto, sino entonces se aprovechaba [para incluir lo] que necesitaba la comunidad. Entonces [se solucionó con] el tema del nexo causal entre los hechos, los daños, y las medidas reparadoras" (Asesora para la formulación de política pública, Subdirección General-UARIV).

> "… me parece que está ligada con esta [segunda con tercera recomendación]. Ósea, primero define primero quienes son sujetos de reparación colectiva, segundo define que tienen que tienen que tener esos sujetos, define los componentes; entonces si la medida no corresponde a un componente material, a un componente político, o a un componente simbólico, no puede ir en el plan (…) Eso también lo define ahí en el decreto, eso está en reforma, el nexo causal, las medidas tienen que reparar el daño que me causo el conflicto como comunidad. Entonces, el ejemplo típico del hospital, ¿qué daño le está reparando el hospital? caso contrario el puente, el puente si tiene que ser reconstruido y es un medida de desarrollo, es una medida de infraestructura que muy costosa, pero el puente tiene que ser reconstruido porque lo dinamito un grupo armado, y aparte es una comunidad a la que solo se llega por el Rio Magdalena y por ese puente, por el Rio Magdalena la moto canoa vale 180 mil, mientras que allá los lideres tienen motos o bicicletas y no las pueden usar porque tumbaron el puente. Si tú les pones el puente, ellos no tienen que pagar 180 mil para salir, y pueden salir por ahí a estudiar, sacar sus productos, pueden pasar ambulancias, todo. Entonces es como el ejemplo, como ambas son medidas de infraestructura, ambas son medidas de desarrollo, ambas son costosas, pero una no tiene nexo causal mientras que la otra sí, eso sí lo tiene la reforma" (Funcionaria de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas, PGN).

Con lo anterior, puede decirse que la segunda y la tercera recomendación fueron implementadas en la misma dirección planteada por la evaluación, por lo que reflejan el uso y la efectividad de esta. La cuarta recomendación, con la cual se esperaba implementar varias de las doce recomendaciones, se trata de elaborar un documento CONPES para el PRC, el cual tenía la posibilidad de ser una

herramienta estratégica para la definición de parámetros acerca de funciones, competencias y compromisos de las entidades en fases previas a los PIRC (CNC, 2016c, p. 19). Con respecto a su elaboración, se respondió de manera negativa: "… digamos que ese no se ha trabajado, es probable que no se haga, en este momento, con un cambio de gobierno, no podríamos saber si hay un interés particular en expedir un CONPES para ese tema tan particular" (Asesora para la formulación de política pública, Subdirección General-UARIV). En efecto, a la fecha no se ha construido un documento CONPES para el PRC, ni se cuenta con un borrador para el mismo, por lo que la recomendación no se utilizó.

En quinto lugar, y a modo de fortalecimiento de la recomendación anterior, se sugirió elaborar un Decreto de Reparación Colectiva, en el cual se defina el PRC, sus objetivos, sus componentes revisados, su articulación con el documento CONPES, un desarrollo de las funciones y competencias de las entidades que participan en el diseño e implementación del PRC, y mecanismos para que entidades no incluidas en el Subcomité de Reparación Colectiva sean incorporadas en tanto adquieran compromisos en los PIRC aprobados (CNC, 2016c, p. 23). En este caso, la reforma al PRC no contempló elaborar un decreto particular para el tema de reparación colectiva, sino que su modificación se plasmó en la Resolución 3143 de 2018, y debería avanzar con la entrada en vigor del Decreto modificatorio del Decreto 1084 de 2015, en donde entrarían otras modificaciones a toda la Política de Víctimas y Restitución de Tierras:

> "… se está trabajando también, en realidad digamos que el Decreto más que uno pensar en que sea un decreto particular de Reparación Colectiva, no. Hay un decreto sectorial que es el Decreto 1084 del 2015 que es la compilación de todos los decretos de las entidades asociadas al sector de inclusión social y reconciliación. Entonces lo que se ha venido trabajando y que ya se tiene la propuesta, es la propuesta de ajustes a ese 1084 para incorporar los ajustes del PRC, no necesariamente está asociado a que haya un decreto particular" (Asesora para la formulación de política pública, Subdirección General-UARIV).

> "… el decreto va a hacerlo, ósea se espera que salga la modificación al 1084, pero el 1084 recoge muchos temas, recoge desde registro, pero yo creo que no importa si es un decreto especifico o es una parte de otro decreto, lo importante es que lo regulen ¿no?. Entonces yo diría que de esas cinco al menos tres están claramente recogidas en la reforma" (Funcionaria de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas, PGN).

En efecto, dentro del Proyecto de Decreto Modificatorio del Decreto 1084 de 2015 y en la Resolución 3143 de 2018, se incluyen los temas que sugiere la recomendación como cambios normativos. El alto grado de probabilidad de que este decreto sea promulgado, permite establecer que la recomendación fue implementada por lo menos a nivel técnico. La sexta recomendación sugirió la elaboración de un manual operativo del PRC, en donde se hiciera explícita la lógica

del programa, sus principales procedimientos, su marco normativo, el desarrollo conceptual con respecto a la reparación colectiva, y lineamientos precisos acerca de la formulación e implementación de los PIRC (CNC, 2016c, p. 25). Acerca de esto se contestó:

> "Lo hay, digamos que en este momento tenemos un documento que incluso se sustentó y se divulgo en el Encuentro Internacional de Reparación Colectiva, y es el modelo de reparación, entonces digamos que ahí están los principales ajustes del programas, no sé si sea considerado como el manual operativo pero si es el primer documento que ya contempla los ajustes, si no sé si sea denominado manual operativo" (Asesora para la formulación de política pública, Subdirección General-UARIV).

> "Ellos deben tener documentos parecidos, quizás no se llame manual operativo, pero eso ya ellos lo tienen, que el desarrollo conceptual de lo que es la reparación colectiva, cadena de valor, todo eso digamos que tendrían que organizarlo en documento que se llame así" (Funcionaria de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas, PGN).

Al respecto, además de la Resolución 3143 de 2018, por medio de la cual se adopta el modelo operativo de Reparación Colectiva, se pudo constar la construcción de tres documentos a partir de los cuales se implementa esta recomendación. Primero, está el documento "Procedimiento para la gestión de la implementación y seguimiento de los PIRC" del Sistema Integrado de Gestión de la UARIV, con fecha del 5 de diciembre de 2017, en donde se realiza un desarrollo conceptual y se aclaran las actividades puntuales con insumos, salidas y responsables de las mismas. En segundo lugar, está el documento "Manual Operativo Ruta de Atención, Asistencia y Reparación Integral a las Víctimas" (UARIV, 2018a) de la UARIV, con fecha del 16 de agosto de 2018, el cual aborda conceptualmente la dimensión colectiva de la ruta integral de reparación, y desarrolla cada una de las fases de la ruta. Por último, está el "Manual metodológico para la formulación de PIRC bajo la estructura de marco lógico" (UARIV, 2018b), con fecha del 9 de octubre de 2018. En este sentido, la sexta recomendación también fue implementada.

La séptima recomendación, consistió en la elaboración de una batería de indicadores de resultados que permita medir qué tanto los SRC están siendo reparados integralmente, y medir el desempeño del PRC en cuanto a metas y objetivos, y de este modo generar información continua, oportuna y confiable para la toma de decisiones, la rendición de cuentas y la asignación de recursos (CNC, 2016c, pp. 25-26). En cuanto a su grado de implementación, en la UARIV se contestó negativamente, y en la PGN se respondió: "No sé, no creo. ¿Es que esta batería de indicadores no fue lo que planteo el gobierno en el Plan Nacional de Desarrollo?, ahí están los indicadores, hay tres indicadores de reparación colectiva. Entonces si yo fuera el gobierno yo te diría que, si hay, están en el Plan Nacional

de Desarrollo" (Funcionaria de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas, PGN).

Evidentemente, la recomendación no está direccionada a tener únicamente los tres indicadores del Plan Nacional de Desarrollo con el que ya se contaba en 2018 al momento de hacer la entrevista (PND "Todos por un nuevo país 2014-2018"), puesto que la evaluación es del 2016. Posterior a la evaluación, sólo se encontró algunos requerimientos generales en la Resolución 3143 de 2018 (art. 18), y en el "Manual metodológico para la formulación de PIRC bajo la estructura de marco lógico" (UARIV, 2018b), dirigidos hacia la necesidad de hacer seguimiento y evaluación a los PIRC, sin embargo, no se encontraron indicadores formulados en dirección a la séptima recomendación, por lo que está no ha sido implementada.

La recomendación octava, consistió en ajustar la composición del Subcomité de Reparación Colectiva a través de una resolución que expida el Comité Ejecutivo (máxima instancia de decisión del SNARIV), en la cual se debe incluir a: Centro Nacional de Memoria Histórica, ICBF, SENA, Unidad de Restitución de Tierras y Ministerio del Interior. Adicionalmente se deben establecer mecanismos flexibles de vinculación y desvinculación de entidades dependiendo de sus compromisos con respecto a la implementación de medidas dentro de los PIRC (CNC, 2016c, p. 29). Al respecto, se contestó:

> "Eso lo acogieron en la reforma al decreto que te digo, quedo un artículo para la parte de fortalecimiento institucional que dice que la conformación de los subcomités de reparación colectiva podrá ser ajustada en cualquier momento. Ellos lo dejaron así, por qué, actualmente decreto 4800 dice conforme a cada comité a, b, c, d. Pero dijeron no vamos a poner taxativamente que entidades deben conformar un subcomité porque cada vez aparecen y desaparecen entidades. Con el acuerdo final desaparecieron algunas y han aparecido muchas otras, entonces cada vez que aparezca una entidad no se puede modificar el decreto, entonces la propuesta es dejarlo abierto, que cada subcomité podrá definir sus integrantes y cambiarlos en el tiempo que estipula conveniente, entonces ahí se soluciona eso" (funcionaria de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas, PGN).

Tal como manifestó la funcionaria, se pudo constatar que en el Proyecto de Decreto Modificatorio del Decreto 1084 de 2015, en su artículo 46 se establece que la UARIV "podrá conformar nuevos subcomités, ajustar los existentes y establecer las entidades participantes en los mismos, con el fin de garantizar la adecuada implementación y según las necesidades de la política pública de prevención, asistencia, atención y reparación integral a las víctimas" (DPS, 2018, p. 19). No obstante, este es un lineamiento genérico para todos los subcomités del SNARIV y aún no ha sido implementado. Puntualmente con respecto a la composición del Subcomité de Reparación Colectiva, en ninguna resolución se encuentra evidencia de que haya sido modificada. Por este motivo, a pesar de la respuesta afirmativa de la entrevistada, la recomendación no puede considerarse como implementada.

En noveno lugar, se recomendó alinear las metas y los objetivos del Subcomité de Reparación Colectiva con las metas y los objetivos del PRC a nivel de resultados, a partir de un proceso de planeación estratégica concertado dentro del mismo Subcomité, entendido como un proceso de alineación de acciones, procesos, recursos y activos de las organizaciones a unos objetivos y metas definidos como prioritarios (CNC, 2016c, p. 29). Esta recomendación implicaba complementar las metas y objetivos de gestión con las metas y objetivos a nivel de resultado. En cuanto a su implementación, se aclaró:

> "Eso si es necesario hacerlo y no se ha hecho, esto es lo que yo te digo que el Subcomité tiene un Plan Operativo Anual en el 2018 que entre otras cosas no se ha aprobado, porque el POA se aprueba en sesión del subcomité técnico, y como no se ha hecho sesión pues no se ha aprobado. De todas maneras, ellos mandaron el POA con observaciones, nosotros lo conocimos, tiene tres acciones. Ósea el subcomité de reparación colectiva que se relaciona con la implementación de un programa tan trascendental como ese, que tienda a la reconstrucción del tejido social, tiene que tener tres acciones. Nosotros lo comentamos, me pareció bastante escueto, te voy a mandar los comentarios, la matriz, lo de la propuesta decreto. Entonces eso no está alineado y como que no lo han tenido en cuenta" (Funcionaria de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas, PGN).

En consonancia, se pudo obtener un oficio de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas dirigido a la Dirección General de la UARIV (oficio: 000943-2017-616681SJPR-LGC, con fecha del 21 de mayo de 2018), en el cual se comenta con respecto al Plan Operativo Anual del 2018 del Subcomité:

> "… se considera que las metas son bajas (…) Para la PGN resulta irrisorio que solo haya 4 metas en el POA del Subcomité Técnico. Teniendo en cuenta el seguimiento realizado consideramos que pueden plantearse más acciones tendientes al avance en la implementación del programa. Sugerimos que se considere como una meta el trabajo de ajustes que se está realizando a algunos PIRC no étnicos (…) llamamos la atención que luego de seis años de implementación de la Ley 1448 y los Decretos Ley Étnicos, se programen pilotos para establecer acciones institucionales sobres sujetos colectivos" (p. 2).

Teniendo en consideración que esta es la situación luego de dos años de tener disponibles los resultados de la evaluación, se puede establecer que la recomendación no ha sido implementada.

La décima recomendación consistió en ajustar la participación de las entidades del SNARIV en la ruta de reparación colectiva, a través de modificar el procedimiento por medio del cual la UARIV convoca a las entidades tanto del nivel nacional como del territorial. Lo más importante de este ajuste fue la obligatoriedad del visto bueno y la concertación previa con entidades que surgen como responsables de la implementación de medidas, para la aprobación de los PIRC (CNC, 2016c, p. 30), al respecto se comentó:

> "… esto es muy difícil reglamentarlo, que participen antes de su aprobación, antes de su aprobación es en la formulación, pero ¿tu como puedes reglamentar eso, si tú no sabes que entidades son necesarias para cada sujeto? En un sujeto tu puedes necesitar al Centro de Memoria Histórica, al Ministerio de Salud y no más, en otro puedes necesitar al Centro de Memoria, al Ministerio de Salud y al Ministerio de Trabajo no más, en otro solo puede necesitarse a la alcaldía porque es un plan muy pequeño, entonces ¿tu como puedes reglamentar eso si va es de acuerdo a la necesidad? Yo creo que la reglamentación es general y eso está en la propuesta de reforma que dice que en la etapa de formulación cada medida debe tener un análisis técnico, jurídico y financiero y hace mucho énfasis en esas palabras, que se debe hacer con las entidades responsables, entonces ahí ya dependiendo el caso se mira quienes son las entidades responsables, pero no se puede limitar a un a, b, c, d, porque como lo haces si son 700 sujetos diferentes" (Funcionaria de la Procuraduría Delegada para la Paz y la Protección de los Derechos de las Víctimas, PGN).

En este orden de ideas, una recomendación orientada hacia la flexibilización de la participación de las entidades del SNARIV en la ruta de reparación colectiva hubiese sido más pertinente. La justificación que se brinda con respecto a la no implementación de esta recomendación está encaminada a subrayar su poca pertinencia teniendo en consideración las particularidades de la gestión del PRC. La recomendación no fue implementada.

Por último, se recomendó estimar el esfuerzo fiscal que debe realizar el estado para implementar el PRC, la cual debía establecer las instancias responsables de su elaboración, los principales asuntos que debe contener, el tiempo de elaboración que conllevara (CNC, 2016c, p. 30). Tanto en la UARIV como en la PGN respondieron afirmativamente en cuanto a su implementación, y señalaron que dentro del Proyecto de Decreto se encuentra el artículo en el que se asigna un tope con el cual se puede hacer dicha estimación (DPS, 2018, p. 17). Esta recomendación fue implementada.

7.5 Conclusiones

El análisis de satisfacción de estándares de calidad refleja que los mayores problemas de esta evaluación se presentaron en el cumplimiento de los estándares correspondientes a reglas de frontera, en efecto, los estándares de participación de los actores relevantes, enfoque asociativo y consideración de evaluación conjunta no fueron satisfechos en este proceso de evaluación, al dejar de incluir un gran número de entidades responsables de la formulación del PRC, y que hacen parte del Subcomité de Reparación Colectiva. Adicional a esto, existieron problemas en los estándares de evaluabilidad y de realización de la evaluación en el plazo y presupuesto asignados. A parte de esto, se cuenta con un cumplimiento de todo el resto de los estándares de calidad en el proceso de evaluación.

En cuanto a los patrones de interacción y a la coordinación de los participantes de las situaciones de acción de diseño y ejecución de la evaluación, a pesar de la no inclusión de muchas entidades del Subcomité de Reparación Colectivo, se tuvo cuidado de que la UARIV, como participante involucrado, fuera también el usuario previsto de las recomendaciones formuladas. Los funcionarios de la UARIV, como usuarios previstos, revelaron una percepción positiva con respecto a la fluidez del trabajo entre el Comité de Seguimiento y el equipo de evaluación. Esto se confirma al encontrar información que reveló el adecuado grado de reactividad del equipo evaluador y del DNP con respecto a las inquietudes y preocupaciones de los usuarios, en diversos temas de la evaluación como en su diseño metodológico, la definición de la muestra, las adaptaciones de la metodología a lo largo de la ejecución de la evaluación y la socialización de los resultados de la evaluación. En contraposición a este balance positivo, no se encontró evidencia acerca de la planificación o realización de algún seguimiento a la implementación de los hallazgos de la evaluación, o de la búsqueda proactiva del uso de la evaluación por parte del equipo evaluador.

A dos años de tener disponibles los resultados de la evaluación al PRC, la UARIV ha implementado seis de las once recomendaciones realizadas por el Centro Nacional de Consultoría. La no implementación de las recomendaciones restantes (con excepción de la construcción de la batería de indicadores de resultado) se explica en el hecho de que requieren decisiones concertadas entre un gran número de entidades que no fueron incluidas dentro de la evaluación, y que, por ende, no tienen apropiación de sus resultados. Por su parte, la recomendación concerniente a la batería de indicadores resulta irrelevante debido a la heterogeneidad de medidas que existe en los PIRC.

Esta evaluación se realizó en un contexto donde hubo interés de reformular el programa y tomar las decisiones pertinentes al respecto, en una estructura organizacional jerárquica. A pesar de que varias recomendaciones quedaron por fuera del poder e interés de los usuarios previstos, la implementación de varias recomendaciones a través de elaboración del proyecto de decreto modificatorio del

Decreto 1084 de 2015, la Resolución 3143 de 2018, y los documentos metodológicos: “Procedimiento para la gestión de la implementación y seguimiento de los PIRC” (UARIV, 2017b), “Manual Operativo Ruta de Atención, Asistencia y Reparación Integral a las Víctimas” (2018a) y “Manual metodológico para la formulación de PIRC bajo la estructura de marco lógico” (2018b), permiten establecer que hubo apertura por parte de los usuarios previstos de cambiar la intervención con base en los resultados de la evaluación.

8. Efectividad de la Evaluación Institucional y de Resultados de la Política de Seguridad Alimentaria y Nutricional.

8.1 Presentación de la Política de Seguridad Alimentaria y Nutricional (PSAN)

En el CONPES 113 de 2008 se define la Seguridad Alimentaria y Nutricional (SAN) como "… la disponibilidad suficiente y estable de alimentos, el acceso y el consumo oportuno y permanente de los mismos en cantidad, calidad e inocuidad por parte de todas las personas, bajo condiciones que permitan su adecuada utilización biológica, para llevar una vida saludable y activa" (p. 3). Por el contrario, la inseguridad alimentaria y nutricional se presenta ante la existencia de limitaciones en la oferta y demanda de alimentos, en el poder adquisitivo, en la calidad de los alimentos, y en la capacidad de absorber los nutrientes que el organismo necesita para su desarrollo óptimo (G/Exponencial, 2015a, p. 15). La SAN es definida a partir de las siguientes dimensiones o ejes:

- La disponibilidad de alimentos: que es la cantidad de alimentos con los que se cuenta a nivel nacional, regional y local. Se relaciona con el suministro suficiente de estos, frente a las necesidades de la población, y depende de la estructura productiva; los sistemas de comercialización internos y externos; factores productivos como tierra, crédito, agua, tecnología y recursos humanos; el ecosistema; políticas de producción y comercio; y, tensiones sociopolíticas (p. 16).
- Acceso: esta dimensión remite a la posibilidad de las personas de alcanzar una alimentación adecuada y sostenible. Está determinada por: el nivel de ingreso de las personas; su condición de vulnerabilidad; factores socio-geográficos; y, precios de los alimentos (p. 16).
- Consumo: remite a los alimentos que seleccionan y consumen las personas. Dentro de sus factores determinantes están: la cultura; patrones y hábitos alimentarios; educación nutricional; información comercial y nutricional; nivel educativo; y la composición familiar (p. 16).

- Aprovechamiento o utilización biológica de los alimentos: remite a cómo y cuánto aprovecha el cuerpo humano los alimentos que consume. Los factores determinantes son: medio ambiente; la salud de las personas; estilos de vida; situación nutricional de la población, la disponibilidad de alimentos; la calidad y acceso a servicios de salud, agua potable y saneamiento básico; y, fuentes de energía (p. 17).
- Calidad e inocuidad: se refiere al conjunto de características de los alimentos que garantizan que sean aptos para su consumo. Dentro de sus determinantes se encuentran: la normatividad; inspección, vigilancia y control; los riesgos biológicos, físicos y químicos; y la manipulación, conservación y preparación de alimentos (p. 17).

Estos ejes/dimensiones y sus respectivos problemas dentro del territorio colombiano se abordaron a partir de las líneas de política pública que se plantearon en el CONPES 113 de 2007:

- Estabilidad en el suministro y desarrollo del mercado agroalimentario.
- Impulso a las formas asociativas y empresariales para la generación de empleo e ingresos que contribuyan a la disponibilidad y acceso a los alimentos.
- Mejoramiento de la capacidad para acceder a los factores productivos a la población vulnerable.
- Garantía de Acceso a los Alimentos.
- Promoción y protección de la salud y la nutrición, y fomento de estilos de vida saludable.
- Mejoramiento de los servicios públicos, saneamiento ambiental y entornos saludables.
- Aseguramiento de la calidad e inocuidad de los Alimentos.
- Desarrollo científico y tecnológico de los 5 ejes de la seguridad alimentaria y nutricional.
- Desarrollo de las capacidades, potencialidades y competencias humanas.

Adicionalmente, la PSAN cuenta con tres herramientas principales para su implementación: la Comisión Intersectorial de Seguridad Alimentaria y Nutricional (CISAN), como estructura institucional intersectorial para la orientación, la coordinación, la planeación, el seguimiento y la evaluación (p. 29); el Plan Nacional de Seguridad Alimentaria y Nutricional 2012-2019 (PSAN), el Observatorio de Seguridad Alimentaria y Nutricional (OSAN); y, los planes territoriales de la SAN.

8.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Seguridad Alimentaria y Nutricional.

8.2.1 Reglas de alcance y resultados potenciales

La documentación resultante de la etapa de diseño de la evaluación, previa a la contratación del equipo evaluador, aclara las razones por las cuales la evaluación se realizó en ese particular momento. Estas razones corresponden a los retos particulares que tiene el país en materia de seguridad alimentaria y nutricional, estos son: sostenibilidad de una oferta estable y saludable de alimentos, acceso a alimentos en las regiones apartadas y por parte de población vulnerable, aumento en la eficiencia del suelo, mitigación del riesgo climático, mejoramiento en la ingesta de proteínas, frutas y verduras, disminución en deficiencias en micronutrientes, disminución en el consumo de alimentos de bajo valor nutricional, lactancia exclusiva, promoción de hábitos y estilos de vida saludable, mejores condiciones de saneamiento básico, fortalecimiento de sistemas de vigilancia y control y condiciones para una adecuada información y orientación a los consumidores (DSEPP-DNP, 2014g, p. 2).

Adicionalmente se justificó la realización de la evaluación a partir de subrayar la importancia que tiene la PSAN a nivel intersectorial, la medición de las metas establecidas en el CONPES 113 de 2007, y la recomendación de este documento de realizar una evaluación quinquenal para la política (DSEPP-DNP, 2014g, p. 3). Así mismo, se aclaró que los resultados de la consultoría deben permitir a las entidades encargadas de la PSAN, utilizar la información para la toma de decisiones y el ajuste de la política (p. 56).

El objetivo general de la evaluación fue: “realizar una evaluación institucional y de resultados de la (…) – PSAN– que permita identificar la capacidad institucional de las entidades a cargo de su formulación e implementación, a nivel nacional y territorial, e identificar el cumplimiento de los objetivos y el avance de las metas, de acuerdo con lo establecido en el documento CONPES 113 de 2008” (G/Exponencial, 2015b, p. 2). Para cumplir con este objetivo general, se establecieron los siguientes objetivos específicos (DSEPP-DNP, 2014g, p. 53):

- “Proponer y desarrollar una metodología de evaluación de la capacidad institucional de las entidades encargadas de la formulación e implementación de la política”.

- “Identificar y caracterizar a las entidades a nivel nacional y territorial y sus factores de éxito o puntos críticos en términos de formulación, implementación y articulación”.
- “Identificar la capacidad del recurso humano de las entidades a nivel nacional y territorial para llevar a cabo la formulación e implementación de la política”.
- “Identificar las fuentes de financiación de la política a nivel nacional y territorial”.
- “Medir el grado de articulación de las diferentes entidades a nivel nacional y territorial que permita identificar las relaciones fuertes y débiles entre actores”.
- “Identificar y analizar el avance de las metas establecidas en el documento CONPES 113 de 2008”.
- “Identificar y analizar los determinantes que explican el logro de estas metas”.
- “Identificar y analizar los resultados de cada una de las líneas de política y su contribución al logro del objetivo de la política”.
- “Identificar cómo la estructura institucional actual incide en la obtención de resultados de la política”.
- “Proponer recomendaciones orientadas a la definición de una estructura institucional de la política que se ajuste a la complejidad e integralidad de las necesidades del país en temas de seguridad alimentaria y nutricional”.
- “Proponer recomendaciones orientadas al cumplimiento de las metas establecidas en el CONPES 113 de 2008”.

El alcance de estos objetivos es definido a partir de los cuatro componentes de la evaluación. El primero es el componente institucional, el cual se dirigió a realizar un análisis de la capacidad institucional de la política a nivel nacional y territorial, abordando dimensiones como el recurso humano, recursos físicos y tecnológicos, asignación de responsabilidades desde la etapa de formulación de la política, la operatividad de la política, y el alcance de esta para abordar los derechos y necesidades relacionados con la SAN (pp. 53-54). El segundo componente corresponde a la evaluación de resultados, en este se abordaron los efectos de la política a nivel nacional y territorial, los factores tratados por la política que son responsables de los resultados (p. 55). El tercer componente brindó lineamiento sobre las conclusiones y las recomendaciones de la evaluación, y el cuarto componente dio lineamientos adicionales para el PTIR, particularmente entorno a la estrategia para la socialización y el uso de los resultados de la consultoría (pp. 55-56).

Todo lo anterior satisfizo los estándares de fundamentación y finalidad, objeto y alcance, y objetivos específicos de la evaluación. En el pliego de condiciones definitivo se establecieron lineamientos para que las conclusiones, recomendaciones y lecciones se formularan de forma clara, pertinente, focalizada y orientada a su puesta en práctica. Esto es abordado por los componentes 3 y 4, los cuales requirieron identificar claramente los aspectos de la política que debieron ser ajustados, desde el cual se construyera un nuevo modelo de arreglo institucional orientado a una mayor articulación y efectividad de los instrumentos de la política (p. 56).

Adicionalmente están los lineamientos específicos del PTIR, en donde destacan lineamientos como la clasificación de las recomendaciones según su capacidad de influencia y facilidad de implementación, el mapeo de actores responsables de las recomendaciones, y la matriz de recomendaciones (pp. 66-67), la cual requirió caracterizarlas con el objetivo de la PSAN a la cual influyen, el objetivo de la evaluación que abordan, el resultado de la evaluación que la justifica, sus mecanismos de implementación, los actores responsables, y los efectos de la implementación. En efecto, todos estos lineamientos fueron cumplidos a cabalidad en el PTIR correspondiente, por lo que se cumplieron con todos los componentes del estándar de oportunidad, pertinencia y utilización que remiten a reglas de alcance.

Por último, en cuanto al estándar de desarrollo de capacidad, no se encontró ningún lineamiento ni objetivo dentro de la documentación del diseño de la evaluación. A pesar de esto, el equipo evaluador estableció como recomendación el fortalecimiento del Observatorio de Seguridad Alimentaria y Nutricional, mediante la financiación de más de una entidad de la CISAN y su orientación hacia un seguimiento y análisis periódico (G/Exponencial, 2015c, p. 32), con esto, se esperaba mejorar la capacidad de monitoreo y evaluación de la PSAN. Esto satisface parcialmente el estándar de desarrollo de capacidad.

8.2.2 Reglas de frontera y participantes

Tanto en el Comité de Diseño como en el Comité de Seguimiento participaron el DNP, el Ministerio de Agricultura y Desarrollo Rural, el Ministerio de Salud y Protección Social y el Departamento para la Prosperidad Social –DPS- (G/Exponencial, 2015a, p. 2), y según las entrevistas, también se incluyó al Instituto Colombiano de Bienestar Familiar –ICBF-. Se debe resaltar que la presidencia de la CISAN es rotativa cada dos años entre el Ministerio de Salud y el Ministerio de Agricultura, mientras que la secretaría técnica de la CISAN se rota entre el Ministerio de Salud y el DPS. En este sentido, estos son los actores más importantes para la política y, por lo tanto, los más relevantes de ser incluidos dentro del proceso de formulación y supervisión de la evaluación.

A pesar de esto, la CISAN es compuesta por otras entidades que no fueron incluidas como actores determinantes de la evaluación, estas son: Ministerio de Comercio, Industria y Turismo, Ministerio de Educación Nacional, Ministerio de Ambiente y Desarrollo Sostenible, Ministerio de Vivienda, Ciudad y Territorio, Instituto Colombiano de Desarrollo Rural, y la Asociación Colombiana de Facultades de Nutrición. En este sentido, los estándares de participación de los actores relevantes, enfoque asociativo y consideración de evaluación conjunta pueden ser descritos como parcialmente satisfechos.

En relación con el equipo de la evaluación, el DNP adelantó un concurso de méritos abierto (DSEPP-DNP, 2014g, p. 5), al igual que en los demás procesos de este tipo, se establecieron requisitos habilitantes jurídicos, técnicos y financieros para la presentación de las propuestas (DNP, 2014d, p. 4). Estas últimas fueron calificadas a partir de los siguientes criterios: desarrollo metodológico, trabajo de campo, estrategia de muestreo, experiencia y publicaciones del equipo de trabajo, experiencia específica del proponente, y apoyo a la industria nacional (pp. 4-5).

Se requirió para el equipo de trabajo un director del proyecto experto en seguridad alimentaria con mínimo maestría y tres años de experiencia; un experto en nutrición con mínimo posgrado y dos años de experiencia; y un experto en sector agropecuario con mínimo posgrado y experiencia de dos años (pp. 15-16). De cuatro firmas consultoras que se presentaron, tres fueron habilitadas, de entre las cuales se seleccionó a la firma G/Exponencial para la realización de la evaluación (Resolución 4220 de 2014 del DNP). Todo lo anterior satisfizo el estándar de equipo de la evaluación.

Finalmente, frente a los recursos de la evaluación, por una parte, se estableció un presupuesto de la evaluación de $785.000.000 que fueron ajustados de acuerdo con la propuesta económica del equipo evaluador (DSEPP-DNP, 2014g, p. 4). Por otro lado, se estableció como responsabilidad específica del equipo evaluador la disposición del equipo de trabajo ofertado en la propuesta, de acuerdo con los requerimientos que se presentaron en el transcurso de la consultoría (p. 3). La verificación de la capacidad financiera de la firma consultora por parte del DNP en el proceso de contratación también garantizó la suficiencia de recursos para el desarrollo de la evaluación. En este sentido, el estándar de recursos de la evaluación fue completamente satisfecho.

8.2.3 Reglas de elección y acciones

La PSAN es una política con un trayecto y un nivel de consolidación considerables. Esto significó que, para la etapa de diseño de la evaluación, los integrantes del Comité de Diseño ya contaban con documentación relativa a diseño de política pública con estrategias y líneas de acción claras, y con un conjunto de indicadores con líneas de base y fuentes de información (Gobierno de Colombia, 2013, pp. 40-45). Al respecto, en el mismo informe de la evaluación se señala:

> “La política de seguridad alimentaria y nutricional tiene un marco lógico exhaustivo, que incluye virtualmente todos los determinantes, inmediatos y remotos, que explican los riesgos de hambre, desnutrición y malnutrición de los colombianos. Este diseño tiene la virtud de la consistencia técnica, que le imprime robustez, a la vez que ofrece la posibilidad de hacer visible que la seguridad alimentaria y nutricional es una responsabilidad multi-sectorial” (G/Exponencial, 2015a, p. 108).

Dentro de los requerimientos técnicos que se le hicieron a la propuesta metodológica se incluyó que los proponentes deberían demostrar un manejo y precisión conceptual con respecto a la PSAN y a la problemática en SAN en Colombia (DSEPP-DNP, 2014g, p. 23), por lo que, desde la fase de contratación de la evaluación, ya se había garantizado la adecuada definición de la intervención y que sus resultados fuesen verificables. El análisis de la viabilidad de la evaluación es complementado con el respectivo análisis de riesgos previsibles que pudieron afectar el equilibrio contractual.

Dentro de los riesgos que se toman en cuenta y a los que se les plantean soluciones y responsables, están (pp. 46-47): el consultor no firma el contrato o se retrasa en la constitución y presentación de las garantías de cumplimiento de contrato; retrasos o incumplimientos en la entrega de los productos; demora en la aprobación de los productos desarrollados por el contratista; o la ausencia de recursos para pagar el valor del contrato. Con este análisis de viabilidad de la evaluación se cumplieron con todos los componentes del estándar de evaluabilidad.

Para la evaluación se realizó un estudio sistemático de planes a nivel nacional y territorial: el CONPES 113 de 2008, el Plan Nacional de Seguridad Alimentaria y Nutricional 2012-2019, en el cual se plantearon 115 intervenciones en su plan de acción (G/Exponencial, 2015b, p. 2), y los planes territoriales de la SAN correspondientes al nivel departamental, y actas disponibles de las reuniones de la CISAN (G/Exponencial, 2015a, p. 24). En este sentido, el estándar de coordinación y alineamiento fue satisfecho.

Al igual que en resto de evaluación, la selección del enfoque metodológico partió de una serie de criterios generales plasmados en el pliego de condiciones definitivo, los cuales se direccionaron tanto a la propuesta metodológica de los proponentes, como al primer producto de la consultoría correspondiente al informe metodológico, en el cual se presentó una metodología perfeccionada con base a los acuerdos realizados en el marco del Comité de Seguimiento.

La metodología correspondió a los dos primeros componentes que delimitan el alcance de la evaluación: evaluación institucional y evaluación de resultados. En cuanto a la primera se requirió hacer una identificación preliminar de las variables que se debían tomar en cuenta y que abordan las preguntas orientadoras de dicho componente, una identificación de las fuentes de información (DSEPP-DNP, 2014g, p. 23), metodología de análisis de la capacidad institucional de la política a nivel nacional y territorial, identificación y caracterización de entidades encargadas de la formulación e implementación de la política, identificación de fortalezas y puntos críticos de la institucionalidad de la política y de su articulación, incluyendo sus roles y relaciones, metodología para una aproximación cuantitativa para medir el grado de articulación (p. 57), análisis de objetivos, actores, flujo de recursos, procesos y actividades con identificación de cuellos de botella que afecten la capacidad institucional para la formulación e implementación, estudio de las razones por las

cuales la capacidad institucional repercute en la provisión de bienes y servicios, y lleva al logro de la política (p. 58). Finalmente, se establecieron como unidades de análisis de este componente a las entidades responsables de la política en el orden nacional y entidades responsables de la política en el orden territorial, también a los Comités municipales de la SAN, Secretarios de Planeación, Secretarios de Salud, Secretarios de Hacienda y a los Consejos Territoriales de Política Social).

En relación con los lineamientos que enmarcan la metodología de la evaluación de resultados, se requirió una metodología que use tanto método cuantitativo como cualitativo, para identificar efectos planeados y no planeados de la política. Puntualmente se requirió el abordaje de: presencia de oferta a nivel territorial, caracterización de la oferta a nivel nacional y territorial, focalización de la oferta a nivel nacional y territorial, diseño institucional de la política, resultados por eje de la política, y el logro de las metas del CONPES 113 (DSEPP-DNP, 2014g, p. 59). Para este componente se establecieron como unidades de análisis a los usuarios finales de cada una de las líneas de política. Estos requerimientos se plasmaron en los distintos productos de la consultoría, particularmente en el informe metodológico y en el informe final. De esta manera, se cumplió con todos los componentes del estándar de selección de enfoque y método.

Se plantearon 22 preguntas orientadoras para el componente institucional, 9 preguntas para el componente de resultados, y 5 preguntas para el componente de conclusiones y recomendaciones, para un total de 36 preguntas orientadoras (pp. 53-56). De estas, se pudieron identificar 14 que corresponden específicamente a alguno de los criterios de evaluación de la OCDE, de pertinencia, eficiencia, eficacia, impacto y sostenibilidad:

Pertinencia:

- ¿Cuál es la capacidad institucional que tiene la política para cumplir con sus objetivos?
- ¿Qué fortalezas, debilidades u obstáculos tiene la institucionalidad para cumplir con la misión de la política?
- ¿Cómo incide la normatividad vigente en la SAN en el alcance de los objetivos de la política?
- ¿Cuál es la capacidad institucional de las entidades territoriales para la implementación de la política de la SAN?
- ¿Existe coherencia entre las acciones establecidas en el plan nacional de seguridad alimentaria y nutricional – PNSAN y los planes territoriales de la SAN?
- ¿La oferta a nivel territorial responde a las necesidades de los grupos vulnerables (pobreza extrema, desplazada por la violencia y los afectados por desastres naturales, población infantil, mujeres gestantes, madres lactantes y adultos mayores y etnias)?
- ¿El diseño de la política, planteado desde el CONPES 113, responde a la complejidad e integralidad de las necesidades del país en temas de seguridad alimentaria y nutricional?

- ¿Cuáles deberían ser los modelos de implementación de la política teniendo en cuenta su capacidad institucional y el contexto de cada territorio?
- ¿Qué cambios se requieren para que mejoren los aspectos de la capacidad institucional, en función del cumplimiento de los objetivos de política de la SAN?

Efectividad:

- ¿La Comisión Intersectorial de Seguridad Alimentaria y Nutricional –CISAN ha cumplido con las funciones asignadas de acuerdo con la ley? En caso de que no se hayan cumplido ¿Cuáles han sido las razones?
- ¿Cuál es el avance de las metas del Conpes 113? En caso de que algunas no se hayan cumplido, ¿Por qué no se han alcanzado? y ¿Qué se debe hacer para lograr su cumplimiento?
- ¿Existen experiencias exitosas en la ejecución de la política a nivel nacional y territorial que puedan ser replicables?

Eficiencia:

- ¿Cuántos recursos se han invertido en la política de la SAN y cuántos usuarios ha atendido?
- ¿Son suficientes los recursos asignados para la política de la SAN? ¿Cuántos recursos se requieren para alcanzar las metas?

Sostenibilidad:

- ¿Las entidades están en capacidad de gestionar recursos y establecer alianzas y redes de apoyo que le permitan consolidar la política?

De todas las 36 preguntas orientadoras, la única que no fue contestada ni en el informe final ni en el PTIR fue: ¿Existe coherencia entre las acciones establecidas en el plan nacional de seguridad alimentaria y nutricional – PNSAN y los planes territoriales de SAN? En este sentido, los estándares de cuestiones a responder, selección y aplicación de criterios de evaluación y respuesta a preguntas evaluativas fueron satisfechos. También se evidencia una adecuada satisfacción del estándar de claridad del análisis, lo que se refleja en el modo de presentación de los hallazgos y conclusiones en el informe final, y en la elaboración de la matriz de recomendaciones en el PTIR (G/Exponencial, 2015c, p. 27). En esta, se evidencia la distinción y la conexión lógica entre objetivos de la evaluación, conclusiones y recomendaciones.

En el resumen ejecutivo, adjunto a la evaluación en el portal de SINERGIA, se presenta la metodología de la evaluación, el mapa de actores, el mapa de la oferta institucional, los hallazgos en capacidad de dirección, capacidad de ejecución, capacidad financiera, capacidad de monitoreo y evaluación, y resultados en

términos de las metas del PSAN. Por su parte, la ficha técnica incluyó una descripción de la política, objetivo de la evaluación, metodología de la evaluación, principales hallazgos, principales conclusiones y recomendaciones, por lo que este satisfizo mucho mejor que el resumen ejecutivo, los componentes del estándar de claridad y representatividad del resumen.

Respecto a las acciones que se debieron plasmar en el informe final de la evaluación (y en el PTIR en este caso), comenzando con el contexto de la intervención, su análisis fue requerido en sus dimensiones normativa, económica y política desde el pliego de condiciones definitivo (p. 58). En efecto, a lo largo del informe final de la evaluación, se evidencia un abordaje sistemático de estas dimensiones, lo que satisfizo el estándar de contexto de la intervención. Así mismo, en el informe final se describió y valoró la lógica de la intervención, para esto, se incluyeron secciones descriptivas para la CISAN, el Plan Nacional de Seguridad Alimentaria y Nutricional (PNSAN); el Observatorio de la SAN y los Planes territoriales de la SAN (pp. 17-21). También se establecieron conclusiones puntuales acerca de la pertinencia de la política, y la consistencia entre esta y el diagnóstico (p. 184), con lo cual se satisfizo el estándar de lógica de la intervención.

En el informe final se realiza una explicación parcial de la metodología de la evaluación. En el caso del componente institucional se incluyó el marco conceptual de los procesos e insumos estudiados (p. 21), las fuentes e instrumentos de recolección de información, y las técnicas de análisis (pp, 24-26). Por el contrario, la metodología de resultados no es descrita en el informe final, únicamente se mencionan algunas abstracciones acerca de las variables que serán medidas: medición de las metas, indicadores e intervenciones, inseguridad alimentaria, prevalencia de subalimentación, monitoreo de precio de los alimentos, y análisis de oferta de estos (p. 123). Es probable que este escaso desarrollo en la descripción de la metodología de evaluación de resultados se deba, por una parte, al amplio consenso que existe en cuanto a los indicadores a medir en materia de la SAN, y por otra parte, a que esta metodología se limitó a comparar indicadores con metas. En todo caso, únicamente se puede considerar como parcialmente satisfecho el estándar de explicación de la metodología de evaluación.

En el informe final se presentan las fuentes de información utilizadas por la evaluación con la profundidad suficiente como para valorar su validez y fiabilidad. Se realizaron 540 encuestas en 78 municipios, 15 entrevistas semiestructuradas a delegados de las entidades participantes en la CISAN, 5 entrevistas a funcionarios de organismos de cooperación internacional, 4 entrevistas a expertos con reconocida trayectoria en temas de la SAN, 15 entrevistas semiestructuradas a responsables de planes departamentales de la SAN, 15 entrevistas a ONGs que trabajan temas de la SAN en municipios, y 3 grupos focales con gremios de productores de alimentos prioritarios, organizaciones indígenas y de campesinos (G/Exponencial, 2015b, pp. 4-5). También se observa un ejercicio de triangulación entre la información cualitativa y cuantitativa a modo de reforzamiento de los argumentos y las conclusiones de la evaluación (G/Exponencial, 2015a, p. 25). Esto

permite establecer que el estándar de validez y fiabilidad de las fuentes de información fue completamente satisfecho.

Adicional a esto, se señalaron algunos de los límites de la metodología utilizada tales como: imposibilidad de obtener la totalidad de las respuestas a los encuestados, por lo que se requirió de una metodología de imputación que ayudara a completar la información con “una serie de valores plausibles con base a información auxiliar” (G/Exponencial, 2015b, p. 4); límites en la precisión de las estimaciones por un coeficiente de variación demasiado alto (G/Exponencial, 2015a, p. 24); y limites en la capacidad argumentativa de la información cualitativa, la cual es descrita como “percepciones subjetivas” (p. 25). Este grado de autocrítica permite establecer que el estándar de mención de cambios y límites de la metodología fue satisfecho. Por el contrario, y al igual que casi todos los casos analizados hasta el momento, no se realizaron menciones acerca de discrepancias entre miembros del equipo evaluador, por lo que este estándar no fue satisfecho.

Puede sostenerse que se consultó como fuente de información a todo el conjunto de actores relevantes. A estos últimos se les garantizó la confidencialidad de sus datos con el fin de garantizar la calidad de la información (p. 25), esto respondió a los requerimientos de confidencialidad que se habían establecido con anterioridad en el pliego de condiciones definitivo (p. 45), en estos, el consultor quedó obligado a utilizar la información de carácter privado o reservado con sujeción a las normas y disposiciones legales vigentes, y a abstenerse de disponer de bases de datos individuales a favor de entidades públicas o privadas. Con esto se cumplió el estándar de consulta y protección de implicados.

Por último, el contrato estipuló que el informe final debió ser entregado el 10 de agosto de 2015, y el PTIR el 10 de septiembre de 2015 (DNP, 2014e, p. 5). En el informe final se aclara que la primera versión del documento fue entregada, en efecto, el día 10 de agosto, aclaración pertinente teniendo en consideración que el documento definitivo que fue difundido en el portal de SINERGIA tiene fecha de octubre de 2015. Por su parte, el PTIR entregado tiene fecha de septiembre de 2015, por lo que ambos documentos fueron entregados cumpliendo con las fechas previstas por el Comité de Diseño de la evaluación. No se encontraron adendas u otrosí que alargaran el contrato o ampliaran el presupuesto de este. Esto satisfizo el estándar de evaluación dentro del plazo y el presupuesto previstos.

8.2.4 Reglas de agregación y control

Las reglas de agregación dentro de este proceso de evaluación se concentraron en el control del DNP del cumplimiento de las obligaciones contractuales del equipo evaluador. El DNP ejerció el control en el cumplimiento del objeto del contrato y fue el encargado de expedir el certificado de cumplimiento a satisfacción (DSEPP-DNP, 2014g, p. 4) de los productos, para posteriormente realizar el pago al equipo evaluador. Se requirió el establecimiento de una garantía (mediante contrato de

seguro en póliza, patrimonio autónomo y garantía bancaria) que cubriera los riesgos de: cumplimiento del contrato, calidad del servicio y pago de salarios (p. 50). El primero y el segundo riesgo cada uno cubierto por un valor de garantía del 30% del valor del contrato, y el tercero por un valor del 10% del valor del contrato.

Al igual que en el resto de los procesos de evaluación analizados, el Comité de Seguimiento funcionó como apoyo técnico a la supervisión del contrato, particularmente realizando un seguimiento a los productos, emitiendo recomendaciones sobre su calidad técnica, y recomendando parámetros de calidad para la ejecución del proceso de evaluación (p. 50). También se constituyó una secretaría técnica que tuvo a su cargo recibir las recomendaciones a los productos de los miembros del Comité de Seguimiento, consolidarlas, depurarlas y darles el tratamiento respectivo (p. 51).

Esta posibilidad de que los integrantes del Comité de Seguimiento realizaran comentarios, y que, con base en estos, los productos fuesen modificados, fue materializada y confirmada a partir de las entrevistas y de información del informe final. En esto último se aclara:

> "El presente documento incluye ajustes realizados de acuerdo con los comentarios del Comité de Seguimiento recibidos por la firma el 20 de agosto, el 15 de septiembre y en la reunión del 9 de octubre de 2015. Las conclusiones y recomendaciones, por disposición contractual, se presentan por separado en el producto 5 de esta consultoría" (p. 10).

Dentro de las obligaciones contractuales del equipo evaluador, se estableció que una vez este recibiera los comentarios a los productos por parte del Comité de Seguimiento, esté contaría con un plazo máximo de una semana para entregar el producto con los ajustes requeridos (DNP, 2014e, p. 4). Únicamente cuando estuviese verificado y aprobado este producto por parte del Comité de Seguimiento, se pudo proceder con el recibo a satisfacción para el pago al equipo evaluador (DSEPP-DNP, 2014g, p. 68). En síntesis, es la misma estructura de control que en los otros procesos de evaluación. Todo esto satisface plenamente los estándares de control de calidad y de inclusión de observaciones de los implicados.

8.2.5 Reglas de posición

Al igual que en el resto de los procesos de evaluación, las reglas de control entraron en tensión con las reglas de posición en el punto de la discrecionalidad que tienen el consultor para la defensa de sus productos y su independencia, teniendo en cuenta el alto costo que le acarrea no satisfacer las exigencias del Comité de Seguimiento. Los siguientes comentarios realizados a los productos del informe final y al PTIR, que fueron incluidos por parte del equipo evaluador, ayudan a dilucidar hasta qué punto hubo un impacto negativo de la estructura de control en las reglas de posición:

> "…fortalecer el análisis de la información recolectada versus el deber ser, con el objetivo de fortalecer el argumento sobre la nueva estructura de la CISAN que proponga el consultor en el producto 5" (DSEPP-DNP, 2015f, p. 3).
>
> "…realizar un análisis sobre el mapa de oferta elaborado, de tal manera que permita informar si la cantidad de entidades y de oferta de programas es: mucha o poca, suficiente o insuficiente, responde a un diagnóstico o no, entre otros" (p. 3).
>
> "…mejorar el análisis sobre Capacidad Institucional ya que el documento se limitaba a presentar estadísticas descriptivas sin darle un valor agregado a la información contenida. Asimismo, se solicitó dar una posición sobre cuál es la capacidad institucional con la que cuenta la PSAN" (p. 3).
>
> "Se recomienda al consultor hacer la sugerencia sobre priorizar otros contaminantes y nutrientes como grasas trans, colorantes artificiales, conservantes, edulcorantes, entre otros" (2015g, p. 4).
>
> "Se aclara al consultor que en el marco de la Mesa Técnica de la CISAN, se ha considerado en varias oportunidades, que para darle una mayor relevancia a la Comisión, esta dependa de la Presidencia de la República. Para ello, se deberá ajustar la normatividad vigente" (p. 4).
>
> "Con respecto a la recomendación No. 14, se consideró que la recomendación debería ir ceñida a la concepción del OSAN como una instancia independiente (…) es más sencillo formular un solo proyecto de inversión (construir el OSAN como una entidad mixta e independiente) que firme convenios con otras entidades de la CISAN para obtener recursos para investigar y hacer seguimiento a la PSAN" (p. 4).

En síntesis, los comentarios que fueron realizados al producto del informe final, por lo general se direccionan a dar mejor sustento a las afirmaciones, y requieren juicios o valoraciones al evaluador sin que estén orientadas a una dirección en particular. Esto no se refleja en los comentarios realizados al PTIR, en donde se direccionó recomendaciones a decisiones previas, y todos estos fueron incluidos por el consultor a pesar de que iban en contra de su independencia. Esto se explica, por lo menos en parte, por la incidencia negativa de las reglas de control en las reglas de posición. En este sentido, los estándares de independencia del evaluador y de proceso evaluativo libre y abierto no fueron satisfechos en este proceso de evaluación.

8.2.6 Reglas de información

En relación con las reglas de información, tal como en la mayoría de los casos analizados, la documentación relativa a la finalidad y expectativa pudo ser fácilmente encontrada en el portal del SECOP. Las entrevistas evidenciaron que el documento de pliego de condiciones definitivo, y sus antecesores (proyecto de pliego de condiciones y estudios previos), fue el resultado de un trabajo conjunto entre los miembros del Comité de diseño de la evaluación. En esta documentación se presenta la finalidad, el alcance, los objetivos de la evaluación, marco general para la metodología, los recursos, el plazo asignado y, en general, todos los requisitos que los proponentes debieron conocer para poder participar en el concurso de méritos abierto, por lo que todos los componentes del estándar de documentación acerca de la finalidad y la expectativa fueron satisfechos.

Con respecto a la necesidad de hacer el informe de la evaluación comprensible, adicional al cumplimiento de los estándares particulares que comprenden reglas de elección, se puede constatar que los requerimientos que van en esta dirección dentro del pliego de condiciones definitivo: “Este informe debe ser de fácil comprensión y contener gráficas, cuadros y otras ayudas visuales que permitan una lectura e interpretación ágil de la información” (DSEPP-DNP, 2014g, p. 65), fueron satisfechos a lo largo de todo el informe de la evaluación. La única falencia consistió en que la presentación de las conclusiones no se presentó de acuerdo con las preguntas orientadoras, tal y como era requerido en el pliego. A pesar de esto, puede considerarse que el documento del informe final y el PTIR tienen la suficiente claridad para ser considerados comprensibles, y por ende, se cumplió el estándar de informe de la evaluación.

Finalmente, encontramos que se le requirió al equipo evaluador entregar los insumos que permitieran una adecuada socialización de los resultados de la evaluación, los cuales debieron estar enfocados en los resultados y las recomendaciones formuladas para mejorar los procesos de toma de decisiones (p. 67). Estos insumos de difusión correspondieron a: la ficha técnica, presentación extendida con explicación exhaustiva de la metodología aplicada, resultados y recomendaciones, presentación simple con una breve descripción de la metodología, el alcance, los resultados y las recomendaciones de la evaluación (p. 68). Estos productos se destinaron a los directores de las entidades que conforman el Comité de Seguimiento a fin de lograr incidencia, analizar su oportunidad y precisar la forma en que los resultados podían ser implementados. Al respecto, en el Ministerio de Salud aclaran:

> “… la difusión de resultados se lleva a cabo en el seno del Comité Técnico de la CISAN y, después, por ejemplo, en nuestro caso especial, G-Exponencial vino y presentó los resultados al interior de la Dirección de Promoción, que es a donde pertenece esta subdirección, y al Viceministro. Ahí se presentan los resultados que tuvimos y cuáles eran los que nos

> competían a nosotros (…) Entonces nos sirve mucho para informar al congreso, los informes que escribe contraloría, procuraduría, defensoría del pueblo, los derechos de petición desde la ciudadanía, desde la academia. Entonces se convierte en un instrumento de entregar información. Y además de entregar información clara y dar los resultados que tenemos, muestra dónde hemos avanzado, dónde estamos bien, dónde no tanto, dónde mejorar. Sin duda es una fuente de información muy buena" (Subdirectora de Salud Nutricional, Alimentos y Bebidas, Ministerio de Salud y Protección Social).

En términos de socialización de la evaluación, se pudo comprobar su difusión interna (en la CISAN y en el Comité de Seguimiento), y externa, en donde adicional a su publicación en el portal de SINERGIA, se puede observar que tiene un público amplio en términos de rendición de cuentas y como fuente de información oficial. En estos términos, el estándar de difusión de la evaluación fue completamente satisfecho.

8.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la Política de Seguridad Alimentaria y Nutricional

Con respecto a las dimensiones seleccionadas para medir la coordinación de los procesos de evaluación, en primer lugar, se puede establecer que muchas de las entidades que pudieron participar dentro del proceso de evaluación como usuarios previstos, no fueron incluidas. Al verificar los destinatarios de las recomendaciones producidas por la evaluación, se encuentran entidades ausentes en los escenarios de coordinación de la evaluación tales como: Ministerio de Ambiente, INVIMA, Instituto Colombiano Agropecuario, Ministerio de Educación, y en general, todo el conjunto de entidades que conforman la CISAN. Esta ausencia impide que estas entidades se apropien de los resultados de la evaluación, reduciendo la posibilidad de que estos sean utilizados. A pesar de esto, las entidades más importantes para la coordinación y la reformulación de la política evaluada, si fueron involucradas en un rol de usuarios previstos:

> "… en esos comités técnicos, donde estaba todo el equipo de G/Exponencial, el coordinador del proyecto, y pues las entidades, que no éramos todas. Éramos las entidades que estamos haciendo el rediseño, que éramos salud, agricultura, DNP, ICBF y nosotros, que somos los que somos más activos dentro de la comisión" (Profesional especializada a cargo de reformulación de PSAN, DPS).

En cuanto a la dinámica que se estableció dentro del grupo de trabajo conjunto (Comité de Seguimiento más el equipo evaluador), las entrevistas reflejan un nivel adecuado de reactividad del DNP y el evaluador frente a las inquietudes y

preocupaciones de los usuarios previstos. La inquietud y demanda de información principal que manifestaron los usuarios previstos fue tener una perspectiva externa acerca de la articulación interinstitucional de la política:

> "¿Qué queríamos ver? El tema de articulación era algo importante, que sabemos que era como el talón de Aquiles. Entonces que fuera evidente y no que lo dijéramos como institución, sino que fuera la evaluación externa que dijera qué falencias teníamos, qué podíamos mejorar y cómo se podía (…) nos mostró temas que para nosotros no eran tan evidentes, como el tema de no tener recursos asignados específicamente para el tema de seguridad alimentaria, y que, tanto en los presupuestos territoriales como en los de la nación, no existe algo que se llame rubro presupuestal de seguridad alimentaria, sino que lo que hacemos las entidades es sumar" (Subdirectora de Salud Nutricional, Alimentos y Bebidas, Ministerio de Salud y Protección Social).

Tal como se pudo observar en los estándares de objeto y alcance de la evaluación, cuestiones a responder y objetivos específicos, se puede constatar que la evaluación se direccionó a responder a esta demanda principal de información. En el mismo sentido, se refleja reactividad al momento en que el equipo de trabajo conjunto determinó los usos de la evaluación, existía un consenso entre las entidades que tuvieron un rol de usuarios previstos, puesto que se concordaba en la necesidad de insumos para la modificación de la política pública:

> "La evaluación de la política se hizo precisamente para que fuera un insumo para hacer la nueva política, y para el nuevo decreto. Entonces, digamos que desde un principio se supo que esa evaluación era el insumo de ese proceso, porque lo que tengo entendido es que los documentos CONPES se hacen, pero llega un tiempo donde tienen que ser evaluados a ver si están siendo efectivos o no. Entonces digamos que la utilidad en sí, pues era saber si era efectiva la política, y a partir de ahí poder definir los ajustes que se iban a hacer a futuro" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Esto se refleja también al momento de determinar las preguntas y la metodología de la evaluación, en donde se puede observar que el DNP y el equipo evaluador, para coordinarse con los usuarios previstos los involucraron escuchando sus demandas y puntos de vista. Al preguntar a las entrevistadas acerca del grado en que se tuvieron en cuenta sus necesidades a la hora de establecer las preguntas orientadoras y metodología, se respondió los siguiente:

> "Yo creo que si [se tomó en cuenta], porque como el diseño de la evaluación se hizo en conjunto con las entidades, y había una mesa de trabajo con todas las entidades donde se diseñó lo que se iba a evaluar, entonces considero que, si se tuvo en cuenta, pues porque fue resultado de un trabajo conjunto. O sea, ya el proceso de evaluación, a pesar de que lo contrató el DNP, se basó en un ejercicio previo que habían realizado todas las entidades (…) lo

> que pasa es que lo que te digo, como la política se trabaja en el marco de la CISAN, entonces el DNP lo que hizo fue hacer una mesa de trabajo con las entidades que hacen parte de la CISAN, y ahí se definió la metodología con la que se iba a hacer la evaluación de la política, y además, pues se definió cómo iba a ser ese proceso de evaluación, ahí se definieron qué áreas íbamos a trabajar, porque la evaluación tiene dos aspectos: la evaluación de las metas, y otra evaluación institucional; y digamos que las generalidades de esa evaluación se definieron en conjunto con las entidades, entonces yo creo que si se tuvieron en cuenta las observaciones de la entidad [DPS]. No fue que nosotros dijéramos: 'queremos que incluyan esto', no. Sino que en las mesas de trabajo se decidió: vamos a incluir esto, vamos a indagar sobre esto" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Y,

> "Digamos que de aquí había una persona delegada (…) se crea un comité evaluador (…) en el que está DNP, que ponen los recursos de la evaluación y lidera el proceso, y están las instituciones de la CISAN. Sé que, como mínimo, debieron estar Ministerio de Agricultura, DPS y nosotros. ICBF también debió haber estado ahí. Y la firma evaluadora. Entonces se crea el comité y debían estar reuniéndose cada quince días o cada mes, mientras que se definía toda la metodología, los instrumentos para producción de información primaria o secundaria, y ya después la presentación del resultado. Entonces, sí hubo un diálogo permanente entre esas tres instancias: DNP, nosotros como comité técnico y la firma evaluadora, era un diálogo permanente en todos los procesos. Desde la definición de la metodología, la recolección de información, el análisis y la presentación de resultados" (Subdirectora de Salud Nutricional, Alimentos y Bebidas, Ministerio de Salud y Protección Social).

Tal como se puede observar, existe una percepción positiva de los usuarios primarios acerca de la coordinación que existió en el Comité de Diseño de la evaluación, y de la reactividad en el mismo en materia de metodología. Posteriormente, en la etapa de ejecución de la evaluación, también se conservó la misma postura abierta y reactiva: "¿cómo se hacía? nos reuníamos, y el consultor presentaba la metodología, en unas sesiones, y entonces las entidades hacíamos diferentes observaciones, y se ajustaba de acuerdo con esas observaciones. Yo no tengo en mi cabeza así algo como que no se tuviera en cuenta y fuera importante, no" (Profesional especializada a cargo de reformulación de PSAN, DPS). Esto refleja, adicionalmente, que también se realizaron ajustes y adaptaciones metodológicas de manera conjunta.

También se observa involucramiento de los usuarios previstos en la interpretación de los resultados de la evaluación. En cuanto a esto, adicional a los comentarios del Comité de Seguimiento al informe final que fueron incluidos por el evaluador (y que se presentaron en el análisis de reglas de posición) las entrevistadas resaltaron las

revisiones que podían hacer las entidades a los productos del equipo evaluador, y la apertura que tuvo este a considerar los comentarios que las entidades les hacían:

> "… [el equipo evaluador] iban mandando unos resultados, se revisaban desde las entidades y en ese comité se ponían las dudas, y los consensos a los que se llegaban. No recuerdo que las entidades hayan rechazado algo, era más para determinar cómo se iba a enfocar, qué hacía falta y qué no estábamos viendo reflejado en la evaluación (…) sí hubo siempre espacios de socialización y retroalimentación. Y luego, obviamente, los documentos de resultados se revisaron desde las entidades. O sea, se le dio el visto bueno al producto como tal" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Y,

> "… [el equipo evaluador] tienen en cuenta las observaciones que hacían las instituciones. Porque puede ser que nuestros estadísticos confirman la información y no estemos de acuerdo con lo que miremos, ellos tienen en cuenta la opinión de las instituciones participantes y no es que no les importe porque ellos no pagan, no. Reconocen que formamos parte del comité y eso también se le pasa a la firma evaluadora" (Subdirectora de Salud Nutricional, Alimentos y Bebidas, Ministerio de Salud y Protección Social).

Como se puede observar, no existieron grandes desacuerdos entre el equipo evaluador y los usuarios primarios, a la vez que estos últimos, por lo general, eran escuchados. Esto se refleja de manera parecida en el involucramiento de los usuarios previstos en la construcción de recomendaciones. Por una parte, en el PTIR se aclara que esa versión recogió los comentarios realizados por los miembros del Comité de Seguimiento, recibidos el 13 de octubre de 2015 (p. 7), y en el concepto técnico realizado al PTIR, a parte de los comentarios que restan independencia que ya fueron presentados, hay muchos otros que se orientaron al fortalecimiento de los argumentos, la inclusión de nuevos temas, o la corrección de determinados puntos (DSEPP-DNP, 2015g, pp. 3-4).

Al ser puntualmente cuestionadas acerca de algún seguimiento realizado por parte del DNP a la implementación de los resultados de la evaluación, la respuesta fue negativa, por lo que no existió tampoco un plan conjunto de seguimiento que permitiera fortalecer el uso, a través de mantener el espacio de coordinación. Tampoco se encontró evidencia de búsqueda proactiva del uso, y es probable que la cláusula de confidencialidad del contrato de consultoría vaya en dirección diametralmente opuesta.

8.4 Análisis de la efectividad de la evaluación de la Política de Seguridad Alimentaria y Nutricional

La evaluación produjo las siguientes recomendaciones. La primera correspondió a priorizar la agenda intersectorial en el trabajo de la CISAN, es decir, concentrar su actividad en la coordinación y puesta en marcha de intervenciones con una naturaleza intersectorial en la que concurren dos o más entidades (G/Exponencial, 2015c, p. 7).

> "Eso es lo más complicado, porque en la CISAN hay una cantidad de entidades, y desafortunadamente (…) hay algunas que todavía no tienen una idea clara de cuál es su participación en el comité de seguridad alimentaria y nutricional, o creen que su participación es en algo muy puntual. Entonces, es complicado, a pesar de los resultados de la evaluación y del trabajo que ha hecho la comisión durante este tiempo, es difícil que la comisión tenga apuestas intersectoriales en favor de la seguridad alimentaria. Lo que hace la comisión en últimas es tomar lo que tiene cada sector y sumarlo, pero no como una apuesta que se desarrolle intersectorialmente (…) Es muy difícil, porque desafortunadamente los sistemas de contratación pública son muy rígidos, los tiempos de contratación de las diferentes entidades también. Muchas veces la focalización de la gente que trabaja acá es distinta. Entonces, esas sinergias que uno quisiera que se dieran no son tan fáciles en la realidad" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Esto significa que no se ha avanzado en materia de coordinación intersectorial a partir de la evaluación, sin embargo, tiene un poco más de complejidad que la sola implementación o no de la recomendación. En el informe final de la evaluación se reconoce que la CISAN, como comisión intersectorial, tiene como objeto "dirigir y coordinar la PSAN, y servir como instancia de concertación entre los diferentes agentes de esta, de armonización de políticas entre los diferentes sectores involucrados" (G/Exponencial, 2015a, p. 17), y es creada mediante el Decreto 2055 de 2009. Esto significa que la priorización que la recomendación sugiere, ya está incluida en la institucionalidad jurídica de la política pública.

El resultado de la evaluación que justifica la recomendación (G/Exponencial, 2015c, p. 27), resalta que la CISAN no tiene incidencia en las tareas intersectoriales, y que las entidades participantes perciben que en la CISAN se trabajan temas que no son de su competencia. Pero estos problemas se relacionan más con el modo en el que funciona la CISAN que con su objeto o priorización. En este sentido, la recomendación sugiere algo que ya existe, es irrelevante al no dirigirse al problema

que pretende solucionar, y, tal como se puede constatar en la entrevista, no se han implementado acciones adicionales en ese sentido.

La segunda recomendación fue construir e implementar las cadenas de valor de intervenciones intersectoriales previstas en el PNSAN, precisando las dependencias y funciones para cada etapa de la cadena y las fuentes de financiación para su implementación (G/Exponencial, 2015c, p. 7). Según la matriz de recomendaciones el mecanismo de implementación fue solicitar a las entidades responsables de las 23 intervenciones inter-sectoriales construir la cadena de valor inter-sectorial para cada una de ellas (p. 27). Al respecto, en el DPS se respondió:

> "Pues no es que hayamos trabajado puntualmente algo que ver, cuando vamos a trabajar las cadenas de valor intersectoriales, por lo que te decía. No es que haya una apuesta intersectorial en la comisión, no es que digan que vamos a sacar una intervención articulada que sume del ministerio de agricultura, de nosotros y que llegue salud. Eso no existe. Entonces no es que nosotros hayamos construido una intervención que permita construir cadenas de valor intersectoriales" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Esta recomendación tampoco fue implementada. La tercera recomendación está relacionada y es parecida a la primera: "Agendar preferentemente temas inter-sectoriales en el temario de la CISAN" (G/Exponencial, 2015c, p. 8), esto a partir de dedicar mayor parte del temario de las reuniones de la CISAN a la discusión de temas intersectoriales. Al preguntar acerca de su grado de implementación se respondió:

> "… una de las cosas que hemos hecho es que, en ese plan de trabajo que tiene la Comisión cada año, intentemos trabajar temas que sean del interés de todos, un poco intersectoriales. Entonces ya no, que era lo que pasaba en alguna época, prosperidad social priorizaba algún tema porque era de su interés. Lo que hemos hecho es, aunque el tema de su interés sea interesante, busquemos temas que sean comunes para que todos trabajemos en pro de él y sea algo más intersectorial. Pero en términos de lo que es la oferta de las entidades para responder al tema de seguridad alimentaria, es mucho más complejo. Pero la comisión sí hace el esfuerzo de tener un plan de trabajo que se aborde intersectorialmente" (Profesional especializada a cargo de reformulación de PSAN, DPS).

A pesar de la similitud que tiene la primera recomendación con la tercera, en esta última si se reflejan avances concretos que permiten establecerla como implementada. La cuarta recomendación sugirió gestionar la implementación de las intervenciones del PNSAN que no registran ningún avance (43 en total). Para esto se propuso solicitar a las entidades responsables incluir en sus planes de acción anuales las actividades necesarias, e identificar con las oficinas de planeación de cada entidad, los recursos necesarios para su implementación (G/Exponencial, 2015c, p. 10). Al preguntar al respecto se contestó:

> "… las intervenciones no tienen una temporalidad igual dentro de las entidades, y eso dificulta, algunas veces, ese trabajo intersectorial como uno quisiera, en la implementación. Es decir, nosotros como CISAN sí nos sentamos, rediseñamos la política, entonces metemos todo lo que, como sectores, como el deber ser, lo que debería ser la política. De ahí a trabajar esa información como uno quisiera, pues hay veces en las que uno sueña, cuando uno piensa en las comisiones intersectoriales, se imagina que ellas toman esa política y hacen intervenciones puntuales para resolver necesidades, o dar línea a las entidades y darles instrucciones justificadas, o darles espacios a intervenir. Eso no pasa mucho. Y no pasa mucho porque la CISAN es una comisión que reúne ministerios y departamentos administrativos que tienen un rango del mismo nivel, entonces no permiten que se den órdenes entre ellos por la igualdad. Entonces, una de las cosas que estamos incorporando en ese rediseño de la política es que la Presidencia de la República debería ser quien coordine ese CISAN" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Nuevamente se señalan los problemas de coordinación intersectorial de la política a modo de justificación por la no implementación de la recomendación. En quinto lugar, se recomendó interpretar los lineamientos de la SAN de la Corte Constitucional en el Plan de Acción, esto mediante el desarrollo de la definición de la SAN que adoptó la Corte, teniendo en cuenta la protección de los procesos de los alimentos al productor y al consumidor (G/Exponencial, 2015c, p. 27).

> "… lo que se hizo fue crear una sub-mesa que se enfoca en los temas étnicos, y ahí se trabaja toda la articulación y todo lo que tiene que ver con lineamientos para responder, no sólo a los autos, sino a los lineamientos de política para población étnica. Entonces, se han incorporado en los planes de trabajo de la comisión dependiendo de lo que se esté solicitando, lo trabaja una mesa, o una mesa técnica. Es, un poco, lo que hemos hecho, incorporarlo un poco a la línea de trabajo (…) Hay muchos que cuestan demasiado y son complejos. Sobre todo en el tema étnico, es lo que más complejo es para las entidades" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Al incluir los lineamientos de la Corte Constitucional, en los planes de trabajo han implementado la recomendación. La sexta recomendación consistió en impulsar el desarrollo del enfoque "De la granja a la mesa", a través de un acuerdo intersectorial entre el Ministerio de Agricultura y Desarrollo Rural, Ministerio de Salud y Protección Social, Ministerio de Ambiente y Desarrollo Sostenible, Ministerio de Comercio, Industria y Turismo, y el sector privado, el cual debe definir insumos, procesos y resultados de cadenas agroalimentarias (G/Exponencial, 2015c, p. 16). Al respecto se respondió:

> “Pues nosotros trabajamos eso como prosperidad social, porque nuestros programas son programas de producción de alimentos para el autoconsumo. Pero tienen el componente adicional de cómo se consumen, cómo se preparan. Estos programas que tienen el enfoque ‘De la granja a la mesa’, nosotros somos el único que hace parte de la CISAN, que impulsamos estos programas de autoconsumo. Y, pues, los hemos impulsado porque Prosperidad Social tenía sólo un programa que nació en 2003 y con el tiempo se han creado otros programas que tienen ese mismo enfoque para diferentes poblaciones. Entonces, ahora tenemos uno para la población que retorna a sus lugares de origen, tenemos la que está dedicada únicamente a población afro e indígena, otro que tuvimos que era para población campesina, y nosotros que es para todo el sector de la población que trabajamos nosotros (…) es algo que nosotros hemos venido haciendo desde antes, incluso, de que estuviera la política. Es un problema que ya existía y que prosperidad social ha impulsado, y no porque la evaluación se lo diga, sino porque nosotros consideramos que el primer eslabón que se debe trabajar con las familias más pobres es el tema de seguridad alimentaria y nutricional. A partir de ahí, se trabaja en generación de ingresos, o lo que sea. Pero que sea una consecuencia de la evaluación, no” (Profesional especializada a cargo de reformulación de PSAN, DPS).

Por un lado, se tiene que la recomendación no ha sido implementada. Si el DPS considera que son los únicos que impulsan este enfoque, es claro que este acuerdo interinstitucional para impulsar este enfoque, no se ha realizado. Adicionalmente, la recomendación corresponde a una justificación de decisión previa a los hallazgos de la evaluación, puesto que el DPS ya estaba impulsando este enfoque, y ya había adelantado acciones en ese sentido, no significó un cambio en el modelo mental de los usuarios previstos, no mejora el proceso de toma de decisiones, y no refleja efectividad del proceso de evaluación, además uno de sus componentes es solicitado como recomendación de manera explícita por parte del Comité de Seguimiento: “Se recomienda al consultor hacer la sugerencia sobre priorizar otros contaminantes y nutrientes como gradas trans, colorantes artificiales, conservantes, edulcorantes, entre otros” (DSEPP-DNP, 2015g, p. 3).

La séptima recomendación fue precisar las entidades y dependencias responsables de la implementación de las intervenciones del PSAN que no tienen responsable (G/Exponencial, 2015c, p. 17), el mecanismo de implementación propuesta consistió en “precisar las dependencias responsables de las intervenciones que ya existen y definir las que hacen falta, en concertación con las entidades involucradas” (p. 29). Al respecto se comentó:

> “Eso se identificó en la evaluación, y se había identificado antes en un proceso de seguimiento que la secretaría técnica, que la tenía el Ministerio de Salud. Había dicho que había unas acciones que no tenían responsable como tal. Realmente creo que no se ha trabajado mucho sobre ese tema (…) Lo que no creo que haya sucedido es que haya alguna entidad que lo haya hecho porque la recomendación se lo dijo, sino que van modificando sus

> programas" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Tal como se observa, el sentido de la recomendación es definir responsables para las intervenciones que tenía el PNSAN 2012-2019. La vigencia de dicho plan estaba por terminarse al momento de realizar la evaluación, a lo que se le suma el cambio de gobierno en 2018. En la entrevista se puede observar que lo que ha sucedido, en contraposición a definir responsables de lo que ya estaba, es el cambio de los programas en la SAN, lo que significa que la recomendación no fue implementada.

La octava recomendación fue tener en cuenta en la ejecución de la política las prioridades en la SAN que identificó la Encuesta Municipal de Capacidad Institucional, que fue realizada como trabajo de recolección de datos en el presente proceso de evaluación. Esto se esperaba realizar a partir de una solicitud a las entidades miembros de la CISAN (G/Exponencial, 2015c, p. 18). La entrevista aclaró frente a este punto:

> "… eso es algo que queremos que se modifique y, en esa nueva estructura, está así. No está todo perdido en esa nueva estructura con la firma del decreto, porque si no sale por ahí, la idea es buscar e incorporarlo en otro plan, el plan de desarrollo o buscar otras alternativas para que se dé. Para efectivamente lograr recoger eso. De hecho, esa nueva estructura que tenemos diseñada, a nivel nacional tiene representantes de los municipios, de los departamentos, además de lo que tiene que ver con participación de sociedad civil que no está (…) lo que hicimos fue incorporar a esa nueva estructura institucional esa recomendación de tomar [en cuenta] lo de los municipios (…) pero nosotros lo hacemos como un acompañamiento técnico territorial, lo que también queremos es que la participación de esos municipios y departamentos sea una participación un poco más activa dentro de la definición de política pública. Porque se supone que la CONASAN, o la CISAN el día de hoy, es la instancia que coordina la política pública" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Al momento de recolectar los datos, la CISAN contaba con un proyecto de decreto[63] en el cual se buscaba abordar varios de los resultados de la evaluación (incluyendo parte de la presente recomendación), y se iba a cambiar el nombre de la CISAN por CONASN. Este proyecto de decreto estaba firmado únicamente por tres de las entidades que conforman la CISAN. Al indagar por las razones por las cuales no se firmó por el resto de las entidades se aclaró:

[63] Este proyecto se puede encontrar en el siguiente enlace: http://www.prosperidadsocial.gov.co/normatividad/Documents/Forms/AllItems.aspx

> "… cuando pasó a los jurídicos, hasta ahí llegó el decreto. Ahí ya estábamos listos, no era sino firmar, y no pasó. Porque siento que estamos en un momento también que es poco estratégico (…) porque al plan ya se [le] acaba la periodicidad en 2019, y va a tocar cambiarlo, junto con el enfoque, el gobierno, los funcionarios" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Si bien se espera que la modificación de la estructura institucional, responsable de incluir en la ejecución de la política las prioridades que se identifiquen en los municipios, sea implementada a través del nuevo plan que se haga a nivel nacional en materia de la SAN (luego de 2019), esto aún no sucede. También es de desatacar que la respuesta que se ha considerado a esta recomendación no incluye implementar la política con base en los resultados de la encuesta realizada por la evaluación, sino que se piensa una estructura organizacional que realice esta función de manera periódica. Por estas dos razones puede argumentarse que la recomendación no fue implementada.

La novena recomendación fue convocar a la CISAN para definir una respuesta frente al comportamiento de las metas críticas evidenciadas en la evaluación (G/Exponencial, 2015c, p. 22). Con respecto al mecanismo de implementación se sugirió: "Definir respuesta frente a metas críticas, en particular lactancia materna, área agrícola cosechada, y sobrepeso de hombres y mujeres" (p. 30). Al respecto se comentó lo siguiente:

> "… convocar a la CISAN para definir una respuesta frente a ese comportamiento, no, eso no es que haya pasado. Como 'estamos en rojo en este tema', no. Es decir, las entidades sí saben que hay unas metas que están más rezagadas. De hecho, con los resultados de la encuesta de situación nutricional del 2015 que apenas está saliendo, evidentemente se ve que hay unos temas que las entidades tienen que trabajar más. Pero no es que haya habido un proceso de que salió la evaluación y lo que dice que hacer lo hacemos, no" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Para que se tome una mejor decisión, es necesario que se revise el modelo mental de la política con base en un cambio en las acciones posibles y en los resultados esperados. En este caso, la recomendación no realiza este cambio, precisamente porque no sugiere nada concreto como respuesta a las metas críticas, únicamente transfiere su responsabilidad de formular una solución a la CISAN, quien ya la tenía desde un principio. Aun así, y a pesar de su grado de simplicidad, la recomendación no ha sido implementada, y corresponde a la justificación de una decisión previa.

La décima recomendación sugiere que se establezca una presidencia intersectorial para la CISAN, es decir, que la presidencia de la estructura interinstitucional de la PSAN quedara a cargo de una entidad que cuente con capacidad técnica de interlocución intersectorial, y que tenga un rol misional no alineado con un eje de política en particular (G/Exponencial, 2015c, p. 22). En cuanto a esto, se respondió:

> "La evaluación decía que tenía que haber una entidad, que podía ser el DNP, que no tiene un interés en un sector particular sino en todos, eso puede hacer que la comisión sea más intersectorial. Después de las discusiones, la CISAN dijo que necesitaba alguien que fuera de un nivel superior, como la Presidencia, porque no sé si tú conoces la comisión de primera infancia, que es coordinada por la Presidencia, y tienen una oficina y gente en la Presidencia, y son los que mueven el tema, que es prioridad para la primera dama (…) El tema de seguridad alimentaria, quisiéramos que fuera igual, que tuviera un protagonismo en Presidencia y fueran ellos los que halasen el tema" (Profesional especializada a cargo de reformulación de PSAN, DPS).

En efecto, se puede constatar que en el proyecto de decreto se quiere establecer a la Presidencia de la República o a su delegado como quien presidirá la CONASAN (que remplazaría a la CISAN). Sin embargo, la implementación de este decreto, tal como se analizó más arriba, se ve en el mejor de los casos, lejana. Tampoco se ha realizado la modificación a la actual CISAN en dicho sentido. Mas significativo aun, es que esta recomendación haya sido solicitada por parte del Comité de Seguimiento como parte de sus comentarios al PTIR, lo que hace que la recomendación sea una justificación de decisión previa:

> "Se aclara al consultor que en el marco de la Mesa Técnica de la CISAN, se ha considerado en varias oportunidades, que para darle una mayor relevancia a la Comisión, esta dependa de la Presidencia de la República. Para ello, se deberá ajustar la normatividad vigente" (DSEPP-DNP, 2015g, p. 4).

La onceava recomendación fue crear una sub-mesa técnica sobre precios de los alimentos al consumidor, la cual debió incluir al DNP, Ministerio de Agricultura, DPS, MinCIT, la Superintendencia de Industria y Comercio, y la Asociación Nacional de Empresarios ANDI (G/Exponencial, 2015c. P. 31). Con respecto a esto, se aclaró:

> "… [esa recomendación] también nos ha servido para el proceso de rediseño. Por ejemplo, la evaluación de la política nos dijo que no trabajábamos sobre el precio de los alimentos. Y uno cae en cuenta. Entonces a la hora de hacer el rediseño que te estoy contando, de decir que vamos a construir una nueva política, evidentemente el precio de los alimentos es algo que estamos teniendo en cuenta en esa nueva política, porque se hizo visible acá entonces hay que meterlo de alguna manera. Obviamente tenemos que incluir acciones, no que trabajen el tema de precios como regulación, porque eso acá en Colombia riñe un poco con otros temas, pero sí a partir del precio de los insumos productivos, etc. Entonces sí hay recomendaciones de la política que hemos incorporado a este rediseño de política pública y a las nuevas estrategias que tenemos en la política" (Profesional especializada a cargo de reformulación de PSAN, DPS).

No se encontró ninguna referencia a la creación de esta subcomisión dentro del proyecto de decreto. Sin embargo, lo más probable es que estos cambios pretendan ser incluidos al interior del nuevo plan en la SAN que se realice posterior a 2019. Aun así, este plan no ha sido formulado, y no se tiene certeza de que esto sea, en verdad, incluido en la PSAN. Adicionalmente, se puede subrayar que no es equivalente el hecho de que se incluyan acciones (que por lo pronto están indeterminadas) en dirección al vacío de la política en relación con los precios de los alimentos, a que se conforme una mesa técnica para su abordaje. Esto significa que a pesar de que es un problema que se sabe requiere solución, la recomendación específica que produjo la evaluación no ha sido implementada.

En doceavo lugar, se recomendó fortalecer el Observatorio en la SAN estableciéndolo como una entidad de naturaleza mixta, asegurando su financiación a partir de un proyecto de inversión formulado desde la Secretaría Técnica de la CISAN, en el cual deben también suscribirse convenios con otras entidades de la CISAN para movilizar recursos (G/Exponencial, 2015c, p. 23). También se sugiere que se oriente la actividad del Observatorio al seguimiento y análisis periódico, y no solo a la generación de indicadores (p. 31). Al indagar sobre su implementación, se comentó:

> “El observatorio tiene una gran dificultad, y es que, digamos que el Ministerio de Salud lo creó durante el tiempo que fue secretaría técnica. Tenía un sistema de seguimiento al PNSAN, y lo creó con recursos del Ministerio de Salud. Entonces, estas plataformas, el observatorio y el seguimiento, quedaron incorporados a las plataformas tecnológicas, o digitales, del Ministerio de Salud. Entonces, no hay otra entidad, las otras entidades no pueden decir que le van a dar dinero para que fortalezca una plataforma que está en otro ministerio. Entonces si el Ministerio de Salud no le pone recursos al observatorio, pues ese observatorio va a seguir perdido. ¿Qué pasa ahora? El ministerio de salud no es presidente de la CISAN, ni secretaría técnica, entonces no tienen cómo justificar que tienen que dejar esos recursos para el observatorio” (Profesional especializada a cargo de reformulación de PSAN, DPS).

En este sentido, la recomendación no pudo ser implementada por dinámicas y procesos inherentes a las mismas entidades encargadas de la CISAN, y a sus relaciones. Más allá de esto, se encontró evidencia de que la recomendación fue solicitada por parte del Comité de Seguimiento, por lo que es una justificación de una decisión previa:

> “…se consideró que la recomendación debería ir ceñida a la concepción del OSAN como una instancia independiente (…) es más sencillo formular un solo proyecto de inversión (constituir el OSAN como una entidad mixta e independiente) que firme convenios con otras entidades de la CISAN para obtener recursos para investigar y hacer seguimiento” (DSEPP-DNP, 2015g, p. 4).

La treceava recomendación fue fortalecer la calidad del seguimiento de la política adoptando la escala ELCSA (Escala Latinoamericana y Caribeña de Seguridad Alimentaria) como parte del seguimiento de la PSAN, aumentar la frecuencia de la Encuesta Nacional de Situación Nutricional, hacer seguimiento a población vulnerable priorizada en la PSAN, modificar indicadores de acceso, entre otros (G/Exponencial, 2015c, p. 24). Al respecto, se responde:

> "Es que el observatorio tiene una plataforma que es el seguimiento y evaluación a los planes, entonces ahí se supone que las entidades van incluyendo esos datos de cómo van sus acciones en SAN. Para no hacerlo en un Excel, hay un sistema para eso y está alojado en el ministerio de salud" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Sin embargo, a pesar de que esta plataforma aun exista y no se utilice, es necesario avanzar en el proceso de reformulación de política antes de determinar los nuevos indicadores con los cuales se le va a hacer seguimiento. Además, como el seguimiento depende del Observatorio de la SAN, tiene los mismos problemas para su implementación que la recomendación anterior. Esta recomendación tampoco fue implementada. En síntesis, se puede decir que han existido serias dificultades para la implementación de los resultados de la evaluación. En el DPS explican del siguiente modo esta situación:

> "Además estamos en un momento muy particular en la CISAN. Creo que salió en 2016, lo que hicimos fue proponer rediseñar la política de acuerdo con las recomendaciones. De hecho diseñamos un nuevo diseño institucional, es decir, una nueva estructura de la CISAN y de sus instancias y cómo debería funcionar a nivel departamental y territorial. Incluimos lo que estaba en el acuerdo de paz. Pero llevamos un tiempo en la CISAN en una especie de letargo, porque, desafortunadamente, ese decreto que soportaba esa estructura y esa nueva institucionalidad que recoge las recomendaciones de la evaluación, porque dice que la secretaría técnica no se debe rotar, debe ser sólo una, esto lo debe coordinar la presidencia, no ha pasado" (Profesional especializada a cargo de reformulación de PSAN, DPS).

Esta falta de liderazgo, sumada la coyuntura de cambio de gobierno, enfoque y funcionarios a cargo de la PSAN, que también se mencionaba en el mismo DPS; remite a problemas en la variable de oportunidad y pertinencia de la evaluación. Por último, es importante subrayar que, la existencia de recomendaciones que constituyen decisiones previas de los usuarios previstos coincide con que la evaluación realmente no amplió o transformó el espectro de acciones o decisiones posibles, por lo menos, esta es la impresión general que se puede recoger de los usuarios previstos:

> “Pues yo creo que la evaluación hizo fue oficializar lo que ya sabíamos las entidades. Por ejemplo, tú ves que la evaluación en el tema de la evaluación institucional, los resultados que nos da la evaluación es algo que ya conocíamos, porque lo trabajamos todo el tiempo y uno se va dando cuenta de qué es lo que está pasando. Y como esa evaluación se hizo con las encuestas que nos hicieron a todos los que participamos en la evaluación, ya sabíamos que esos iban a ser los resultados. Lo que obtuvimos fue que estaban oficializados, ya no lo dice una persona, sino que están ahí. Pero, no fue algo que fuera increíble. Por lo menos en ese tema institucional, lo que hizo fue darnos la razón de lo que ya veníamos hablando” (Profesional especializada a cargo de reformulación de PSAN, DPS).

Las entrevistas realizadas no puntualizan, en la mayoría de los casos, si las recomendaciones coinciden con acciones que ya se habían considerado necesarias, o que se pensaba implementar antes de tener los resultados de la evaluación. Sin embargo, este fragmento de la entrevista, junto con el grado de obviedad y simplicidad de las recomendaciones frente a los problemas que los usuarios previstos manifestaron en su rol de fuente de información, hace pensar que muchas de estas están formuladas en dirección a justificar decisiones que surgieron del nivel técnico, pero que encontraron en el nivel directivo fuertes obstáculos. Si bien, no se puede decir que se tiene evidencia concluyente al respecto, si es por lo menos un fuerte indicio de justificación extendida de decisiones previas.

8.5 Conclusiones

El análisis de aplicación de estándares profesionales de calidad refleja que sólo existe incumplimiento total del estándar de menciones acerca de discrepancias entre miembros del equipo evaluador. Los otros que tienen problemas no constituyen un incumplimiento total, sino parcial. Estos estándares fueron: desarrollo de capacidad, participación de los actores relevantes, enfoque asociativo, consideración de evaluación conjunta, explicación de la metodología de evaluación, independencia del evaluador y de proceso evaluativo libe y abierto.

La posibilidad de que los resultados de la evaluación fueran utilizados se redujo ante la ausencia de muchas de las entidades de la CISAN en el equipo de trabajo conjunto de la evaluación. Sin embargo, en este se refleja reactividad del DNP y del equipo evaluador frente a las inquietudes y preocupación de los usuarios previstos, en temas como la determinación de los usos de la evaluación, preguntas orientadoras, metodología de la evaluación, y adaptación de la metodología a lo largo de la ejecución del proceso de evaluación. Adicionalmente, se observa un adecuado grado de involucramiento en la interpretación de los hallazgos de la evaluación.

Las entrevistas no reflejaron un involucramiento de los usuarios previstos en la construcción de las recomendaciones, a pesar de que se evidencian cambios al PTIR sugeridos por el Comité de Seguimiento. Es probable que esta perspectiva de la persona indagada en el DPS al respecto se deba al elevado grado de concordancia que tuvieron los usuarios previstos con los resultados finales de la evaluación (lo cual es señalado por estos), y que es probable que llegara al extremo de no producir recomendaciones que implicaran una ampliación o cambio en las acciones y resultados posibles en las situaciones de acción de la política pública evaluada.

De trece recomendaciones producidas por la evaluación, siete no fueron implementadas, dos fueron implementadas, y cuatro reflejaron evidencia de ser justificación de una decisión previa, además de que ser acciones que tampoco se han implementado. El escaso grado de implementación de estas recomendaciones está relacionado con problemas en la oportunidad y pertinencia de las recomendaciones, particularmente por el cambio de gobierno y la ausencia de liderazgo por parte de la Presidencia de la República. También existen fuertes indicios de que, en general, muchas de las recomendaciones son justificación de decisiones previas que surgieron a nivel técnico dentro de las entidades de la CISAN, pero que, a nivel directivo, terminaron siendo descartadas.

Al observar el grado de apertura hacia cambios institucionales de los funcionarios de la política evaluada, se observa que los funcionarios del nivel técnico estaban dispuestos a esto, e incluso manifestaron que el proyecto de decreto fue el resultado de un trabajo “larguísimo y bastante complejo” del nivel técnico de los sectores implicados, así mismo, se resaltó que la evaluación coincidió con una etapa de rediseño de la PSAN. Sin embargo, los mismos funcionarios identifican como causas de su rechazo el estar en un “momento poco estratégico”, en el que la institucionalización de estos cambios se vio frenada por la implicación de tener cerca un cambio de gobierno central. En este sentido, a pesar de que hubo apertura del nivel técnico a modificaciones institucionales en la política, esta no pudo ser trasmitida al nivel directivo de los sectores implicados, por lo que, en conjunto, no hubo una buena apertura de los funcionarios a cargo de la política a estos cambios.

9. Efectividad de la Evaluación Institucional y de Resultados de las medidas de defensa comercial (salvaguardia y derecho Antidumping)

9.1 Presentación de las medidas de defensa comercial de salvaguardia y derecho antidumping.

El dumping se refiere a la práctica comercial desleal en la cual una empresa exporta un producto al extranjero a un precio inferior al que se fija normalmente en el mercado del país de origen (G/Exponencial, 2016a, p. 2). Esta actividad no está propiamente penalizada dentro de las normas de la Organización Mundial de Comercio (OMC), sin embargo, se han establecido algunos lineamientos en el artículo IV del Acuerdo General sobre Aranceles Aduaneros y Comercio (GATT por sus siglas en inglés) y en el Acuerdo Antidumping de la OMC, los cuales se dirigen a regular derechos antidumping específicos a las importaciones de una determinada procedencia (p. 2).

El término dumping hace referencia a la relación entre precios bajos y altos volúmenes de importaciones que se establecía en los mercados denominados como “dumping ground” (o vertedero) de los excedentes de otro mercado, por lo que remite a una discriminación de precios entre mercados (2016b, p. 22). Adicionalmente, implica la búsqueda de poder monopólico a través de la eliminación de la competencia haciendo uso de estrategias de precios diferenciados en el mercado de exportación.

La práctica de dumping al estar dirigida hacia objetivos monopólicos, por lo general tiene efectos predatorios sobre los sectores industriales de los países receptores de las exportaciones con precios diferenciados (p. 23). Para contrarrestar este problema, las medidas antidumping (delimitadas institucionalmente por la OMC) se orientan hacia la neutralización del daño y a la nivelación de los precios de las mercancías entre el mercado interno y su mercado de exportación, mas no contemplan, en ningún caso, la prohibición del dumping.

En contraposición a las medidas antidumping, que surgen como respuesta a prácticas comerciales desleales, el mecanismo de salvaguardia surge del interés del sistema de comercio multilateral por compartir tanto los beneficios como los perjuicios de la liberalización comercial (p. 24). Ese equilibrio se busca a partir de que todos los miembros asuman los costos cuando un país sufre un incremento abrupto de las importaciones y debe ajustar su industria a esta situación.

Las salvaguardias ayudan a apartarse de forma regulada y temporal de las obligaciones de un país (como la Consolidación de Aranceles o la eliminación de restricciones a importaciones), para así proteger su industria brindándole el tiempo necesario para que se adapte a la nueva coyuntura comercial. Puntualmente, para prevenir o reparar un daño y facilitar el ajuste, se aumenta el arancel o se adoptan restricciones cuantitativas (p. 30).

Estas medidas tienen su origen con la suscripción del GATT en 1947, sin embargo, es hasta los años 90', en concomitancia con el proceso de liberalización comercial, que en Colombia se empiezan a expedir las primeras normas legales que incluyen estos mecanismos de defensa comercial (p. 25). Con relación al ámbito internacional, se tiene el Artículo VI del GATT de 1994, el cual autoriza a los países a adoptar medidas antidumping; el Acuerdo Antidumping de la OMC de 1994; el Artículo XIX del GATT de 1994 sobre salvaguardias; y el Acuerdo de la OMC sobre Salvaguardias. Estas normas fueron adoptadas en Colombia mediante la Ley 170 de 1994.

Según el Decreto 152 de 1998, se puede adoptar una medida de salvaguardia si se ha determinado que las importaciones de un producto han aumentado a tal punto, que causan o amenazan causar un daño grave a una particular rama de la producción nacional de bienes similares, o directamente competidores del producto importado (art. 2), se aplican a las importaciones de productos originarios de otros miembros de la OMC, y se aplican a la totalidad de las importaciones del bien investigado (art. 4). Estas disposiciones son desarrolladas en el Decreto 1407 de 1999, por el cual se establece un procedimiento especial para aplicar una medida de salvaguardia. Por otro lado, los Decreto 2550 de 2010 y 1750 de 2015 regulan la aplicación actual de derechos antidumping.

9.2 Caracterización de las situaciones de acción del proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping.

9.2.1 Reglas de alcance y resultados potenciales

Con respecto a los resultados potenciales de la situación de acción de ejecución de la evaluación, se estableció que la evaluación de operaciones debía ir dirigida a identificar las fortalezas y debilidades de las actividades que se ejecutan en el proceso de investigación e implementación de las medidas de defensa comercial de salvaguardia y antidumping, con la finalidad de generar recomendaciones que permitan mejorar la calidad y la pertinencia del proceso (DSEPP-DNP, 2015h, p. 1). También se esperó que estas recomendaciones mejoraran la estructura institucional del proceso, y que esto redundara en el cumplimiento de los objetivos planteados por las medidas de defensa comercial (p. 3). Esta evaluación tiene un alcance relativamente pequeño en comparación con los casos estudiados hasta el momento, únicamente tiene dos componentes.

El primer componente de la evaluación correspondió a la evaluación de operaciones, el cual pretendió analizar y valorar el proceso por el cual se toma la decisión de adoptar o no una medida de defensa comercial de salvaguardia o antidumping, abarcando aspectos como el mapa de procesos, los productos resultantes, el esquema institucional, los mecanismos de seguimiento, la normatividad que enmarcan dichas medidas, la percepción de quienes ejecutan el proceso y toman la decisión, la percepción de quienes realizan las solicitudes de las medidas, y de las partes afectadas por la imposición de estas (p. 51).

El segundo componente de la evaluación estuvo dirigido a dar lineamientos para la construcción de las recomendaciones de la evaluación. Puntualmente, consistió en realizar un análisis de la información obtenida en el primer componente, de tal manera que se identificaran los aspectos de la operación de las medidas de salvaguardia y antidumping que deben ajustarse para mejorar su proceso de implementación, y garantizar una estructura institucional pertinente para cumplir sus objetivos (DSEPP-DNP, 2015h, p. 53). Se requirió que las recomendaciones de este componente estuviesen soportadas en información cualitativa y cuantitativa. Los objetivos específicos fueron coherentes con la definición del alcance de la consultoría. Estos son: (p. 51):

1. “Evaluar el funcionamiento de los procesos (actividades, metodologías, insumos, entre otros que identifique la firma consultora como relevante) bajo

los que se realiza la investigación y se toma la decisión sobre la adopción o no de una medida de defensa comercial (salvaguardia y derecho antidumping)".

2. "Valorar la estructura y calidad del producto principal de la investigación sobre la adopción o no de una medida de defensa comercial -informe técnico de la investigación- elaborado por la subdirección de prácticas comerciales del Ministerio de Comercio, Industria y Turismo (MinCIT)".
3. "Revisar y analizar la normatividad actual que rige las medidas de defensa comercial en Colombia".
4. "Realizar un análisis sobre la eficacia y conveniencia de las instancias en las que se evalúa y se recomienda al Gobierno Nacional la implementación o no de una medida de defensa comercial".
5. "Evaluar los mecanismos de seguimiento a las medidas de defensa comercial adoptadas e identificar las fortalezas, dificultades y oportunidades de mejora en su implementación".
6. "Indagar sobre las consecuencias de la implementación de una medida de defensa comercial sobre los flujos de comercio, la firma solicitante, el sector al que pertenece y otros sectores que se pueden ver afectados".
7. "Plantear recomendaciones sobre los procesos, las actividades y los productos que permitan fortalecer las investigaciones y la aplicación de las medidas de defensa comercial".

Los primeros cinco objetivos remiten claramente al primer componente de la evaluación. El séptimo a el segundo componente. En el único al que podría argumentarse problemas de coherencia es en el sexto objetivo específico, esto a razón de que se dirige a observar resultados o impactos de las medidas de defensa comercial, lo que se saldría del alcance de una evaluación de operaciones. A pesar de este desajuste, se puede establecer que los estándares de fundamentación y finalidad de la evaluación, objeto y alcance de la evaluación, y objetivos específicos fueron satisfechos en este proceso de evaluación.

Adicional al segundo componente, se establecieron otras reglas en el pliego de condiciones definitivo, encaminadas a la claridad, pertinencia, focalización y orientación al uso de las recomendaciones de la evaluación. En este caso, el PTIR correspondió al cuarto producto de la consultoría, para el cual se requirió incluir los factores a tener en cuenta para hacer uso de las recomendaciones en los procesos de toma de decisión, la metodología de identificación y priorización de recomendaciones, la relación entre recomendaciones y objetivos de las medidas, la relación entre la evidencia de la evaluación y las recomendaciones realizadas, las razones que justifican las recomendaciones, la identificación de actores que afectaran las recomendaciones, y se requirió realizar una mesa de expertos para validar las mismas (p. 60). En consonancia, en el PTIR de este proceso se pudo constatar el cumplimiento de todos estos requerimientos.

Con lo anterior, los componentes del estándar de oportunidad, pertinencia y utilización de la evaluación que corresponden a reglas de alcance (los otros componentes se solapan con el estándar de difusión de la evaluación, o son abordados en el análisis de coordinación), fueron satisfechos dentro del proceso de evaluación. Por otra parte, en el desarrollo del quinto objetivo específico, en el informe final de la evaluación se construye un capítulo específico de evaluación de la etapa de seguimiento a las medidas de defensa comercial (G/Exponencial, 2016b, pp. 176- 187), por lo que el estándar de desarrollo de capacidad también fue adecuadamente direccionado por la evaluación.

9.2.2 Reglas de frontera y participantes

Los participantes de la situación de acción de diseño de la evaluación fueron la Dirección de Seguimiento y Evaluación de Políticas Públicas (DSEPP), la Dirección de Desarrollo empresarial por parte del DNP, y la Subdirección de Prácticas Comerciales del Ministerio de Comercio, Industria y Turismo (MinCIT). En la situación de acción de ejecución de la evaluación fueron estos mismos integrantes (G/Exponencial, 2016b, p. 2) que conformaron el Comité de Seguimiento, más el equipo evaluador de G/Exponencial.

Algunos de los actores relevantes que estuvieron ausentes como participantes en estas situaciones de acción fueron: la Dirección de Impuestos y Aduanas Nacionales (DIAN), quien tiene una fuerte participación en la implementación de las medidas de defensa comercial (p. 48); y los miembros del Consejo Superior de Comercio Exterior: Presidencia de la República, Ministerio de Relaciones Exteriores, Ministerio de Hacienda y Crédito Público, Ministerio de Agricultura y Desarrollo Rural, Ministerio de Minas y Energía, Ministerio de Transporte, Ministerio de Ambiente y Desarrollo Sostenible, Ministerio de Vivienda, Ciudad y Territorio, y el Banco de la República (p. 52). Esto significa que existieron vacíos en los estándares de participación de los actores relevantes, consideración de evaluación conjunta, y enfoque asociativo.

En relación con el equipo evaluador, se pudo constatar que se adelantó un procedimiento de contratación abierto y transparente para su selección. De acuerdo con los documentos encontrados en el portal del SECOP, se adelantó un concurso de méritos abierto (DSEPP-DNP, 2015h, p. 6). El 26 de abril de 2016 se publicó la Resolución 1768 de 2016, por medio de la cual se ordenó la apertura del concurso de méritos abierto, el 17 de mayo se cerró el proceso de selección, en el cual se presentaros únicamente dos propuestas: G/Exponencial Consultores SAS y Fundación Foro Cívico Escuela de Democracia, Derechos Humanos y Participación Ciudadana (DNP, 2016c, p. 2). El 3 de junio de 2016 se adjudica el contrato a G/Exponencial, mediante la Resolución 2135 de 2016.

Dentro de los requisitos habilitantes, se les requirió a los oferentes un equipo de trabajo con los siguientes perfiles: un director del proyecto con título de posgrado

mínimo de nivel de maestría en Ciencia Política, Ciencias Sociales, Derecho, Administración o Economía, y con experiencia mínima de tres años en dirección, coordinación, implementación, evaluación o consultorías relacionadas con economía internacional o política comercial internacional; un experto en política comercial con título de posgrado mínimo de nivel de maestría en las mismas áreas que el director, y con experiencia profesional mínima de dos años en proyectos, programas o investigación relacionados con política comercial; y un experto en procesos, con título de posgrado mínimo de nivel de maestría en ingeniería, administración o economía, y experiencia profesional mínima de dos años en caracterización de procesos en entidades del sector público o privado (DSEPP-DNP, 2015h, pp. 17-18).

En este proceso, al igual que en los demás concursos de méritos abiertos, se utilizaron los requisitos habilitantes jurídicos, técnicos y financieros. De igual manera, una vez se pasaba por la fase de habilitación, se procedió a calificar las propuestas de los oferentes con base en criterios técnicos. Dentro de estos criterios se incluyó: desarrollo metodológico de la propuesta, definición del trabajo de campo, experiencia y publicaciones del equipo de evaluación ofertado, experiencia de este, y apoyo a la industria nacional (p. 26). Con lo anterior, todos los componentes del estándar de equipo de la evaluación fueron satisfechos completamente. Al igual que en los anteriores procesos de evaluación, al interior del pliego de condiciones y del contrato se establecen los lineamientos necesarios para garantizar la suficiencia de recursos para la ejecución de la evaluación, por lo que este estándar también fue satisfecho.

9.2.3 Reglas de elección y acciones

Con respecto a las acciones que deben desarrollar los participantes, se puede constatar que los integrantes del Comité de Diseño realizaron un análisis de la viabilidad de la evaluación. Esto se ve reflejado en la caracterización de la necesidad que se encuentra en el pliego de condiciones definitivo (pp. 2.3), en la cual, se justifica la realización de la evaluación por: la necesidad de contar con un diagnóstico de las actividades, metodologías, y productos del proceso con el que se toma la decisión de adoptar las medidas de defensa comercial; la necesidad de identificar oportunidades de mejora en términos operativos, y de la calidad del análisis que se lleva a cabo para determinar la necesidad de aplicación de una medida; la necesidad de aclarar la pertinencia de las instancias de recomendación al gobierno nacional sobre la adopción o no de una medida; y el interés de generar recomendaciones que permitan implementar mejorar para cumplir con los objetivos de las medidas.

Es de destacar que las medidas de salvaguardia y antidumping, así como sus procedimientos, son determinados a profundidad en los decretos 152 de 1998, 1407 de 1999, 2550 de 2010, y 1750 de 2015, por lo que el Comité de Diseño contó con una definición adecuada de la intervención desde el inicio de la evaluación. Así

mismo, los resultados de estas medidas, que incluyen principalmente la investigación que realiza MinCIT (con la cual se sustenta la decisión de adoptar o no una medida de defensa comercial), los ajustes arancelarios o cuantitativos con respecto a volumen de importaciones, o la recuperación de un sector productivo nacional, pueden ser fácilmente descritos como verificables. Esto significó que se satisfizo el estándar de evaluabilidad.

Al restringirse la evaluación únicamente a las medidas de salvaguardia y antidumping, no fue necesario que el proceso de evaluación se tuviera en cuenta y se coordinara con planes regionales o locales. Sin embargo, la evaluación si tuvo en cuenta el Plan Nacional de Desarrollo 2014-2018, en el cual se estableció la internacionalización de los sectores productivos de bienes y servicios, como un medio dentro de la estrategia de incremento de la productividad de las empresas a partir de la sofisticación del aparato productivo (DSEPP-DNP, 2015h, p. 2). En este escenario, las medidas de defensa comercial se observan como un instrumento de desarrollo productivo, siempre y cuando estén acompañadas de compromisos por parte de los solicitantes, que aseguren que estos se adaptaran a las nuevas condiciones de competencia. Con esto, el estándar de coordinación y alineamiento fue satisfecho.

Frente a la selección del enfoque y método, en primer lugar, se le requirió a los proponentes en el proceso de selección del equipo evaluador, la especificación de la metodología de análisis, incluyendo las técnicas a utilizar, los supuestos subyacentes, la pertinencia de las técnicas, y los principales desafíos para: análisis de proceso que desarrolla la investigación e implementación de las medidas de salvaguardia y antidumping; análisis de roles y competencias de actores involucrados; análisis de calidad y pertinencia de las investigaciones realizadas para determinar la necesidad o no de una medida; análisis del marco normativo de las medidas; análisis de los efectos de la implementación de las medidas sobre los flujos de comercio; identificación de países que sirven como referencia de buenas prácticas; identificación de variables de análisis para responder a las preguntas orientadoras; y la identificación de fuentes de información primaria y secundaria para esas variables (DSEPP-DNP, 2015h, p. 27). También se delimitaron puntualmente los actores que debieron ser tenidos en cuenta como fuentes de información (p. 29).

En segundo lugar, dentro de las responsabilidades del equipo evaluador seleccionado, se requirió que la metodología fuera perfeccionada en aspectos como: la estrategia para generar conclusiones y recomendaciones; instrumentos y fuentes para recolectar información; análisis de la percepción de los actores que intervienen en las diferentes etapas del proceso de evaluación; marco conceptual y teórico; identificación clara de la estructura operacional de los instrumentos de defensa comercial; mapa de procesos de los instrumentos de defensa comercial; valoración de los procesos tomando en cuenta la percepción de los involucrados; relación entre preguntas orientadoras y variables de análisis identificadas (p. 55); indicadores de medición; matriz de consistencia de la evaluación en la que se relacionen objetivos específicos, preguntas orientadoras, variables de análisis,

indicadores y preguntas de los instrumentos de recolección de información; y el cronograma de la evaluación (p. 58).

En consideración de lo anterior, se puede decir que el evaluador tuvo un grado relativo de discrecionalidad en cuanto a la selección del enfoque y método. También se requirió como primer producto de la consultoría un informe metodológico que incluyera todos los aspectos de perfeccionamiento enunciados anteriormente. Con todo esto, lo cual se puede verificar en el informe final de la evaluación, se puede decir que el estándar de selección de enfoque y método fue satisfecho.

Dentro de este proceso de evaluación se formularon un total de 45 preguntas orientadoras, siendo en su gran mayoría, preguntas de tipo descriptivo (DSEPP-DNP, 2015h, pp. 52-53). Dentro de estas, y en correspondencia con el tipo de evaluación que se realizó, se utilizó principalmente el criterio de evaluación de pertinencia. Sin embargo, también se incluyeron preguntas con criterios de evaluación de efectividad y de impacto:

Pertinencia:

- "¿Hay coherencia entre las definiciones, el ámbito de aplicación y los procesos de investigación que se establecen en cada norma que rige las medidas de salvaguardia y derechos antidumping?"
- "¿Cuál es la estructura institucional para la adopción o no de cada medida de defensa comercial? ¿qué fortalezas y debilidades tiene esta estructura?"
- "¿Cuáles son las fortalezas y debilidades de que la autoridad investigadora sea el Ministerio de Comercio, Industria y Turismo?"
- "¿Las instancias en las que se evalúa la adopción o no de una medida de defensa comercial: Comité de Asuntos Aduaneros, Arancelarios y de Comercio Exterior, Consejo superior de Comercio Exterior y Comité de Prácticas Comerciales son adecuadas para garantizar una recomendación pertinente y oportuna?, de no ser así, ¿cuáles deberían ser?"
- "¿Los mecanismos de seguimiento son suficientes para asegurar que la medida tenga un carácter temporal?"

Efectividad:

- "¿la imposición de la medida les permitió a las firmas solicitantes y al sector en general resolver su situación de mercado y adaptarse a las nuevas condiciones de competencia?"
- "¿Qué cambios se requieren en los procesos para mejorar el cumplimiento de los objetivos de las medidas?"

Impacto:

- “Desde la perspectiva de las firmas ¿cómo se han visto afectadas las partes interesadas de la investigación (importadores, demandantes de los productos objeto de la medida, entre otros) ante la imposición de una medida de salvaguardia o derecho antidumping?”
- “¿Cuáles han sido las consecuencias de la implementación de una salvaguardia o derecho antidumping sobre los flujos de comercio? ¿las medidas han sido efectivas en términos de cumplir con su objetivo?”

En consideración con las acciones de los participantes en la situación de acción de ejecución de la evaluación, al verificar los productos 3 y 4 de la consultoría, correspondientes al informe final y al PTIR, se pudo constatar que todas las preguntas orientadoras establecidas por el Comité de Diseño en el pliego de condiciones definitivo fueron contestadas en ambos productos. Con lo anterior, se satisficieron todos los componentes de los estándares de cuestiones a responder, selección y aplicación de criterios de evaluación y respuesta a preguntas evaluativas.

Respecto a la claridad del análisis existe evidencia contrapuesta. Adicional al requerimiento de la matriz de consistencia de la evaluación descrito anteriormente, en el PTIR se presenta cada una de las recomendaciones, de tal manera que son relacionadas con las conclusiones y hallazgos que las sustentan. Por parte del informe final, este contiene cuatro capítulos correspondientes a cada uno de los procesos a evaluar de la operación de las medidas de defensa comercial (previo a la solicitud, investigación, toma de decisión y seguimiento), los cuales están subdivididos en 6 subtítulos correspondientes a seis variables seleccionadas: dirección, asignación, tecnología, logística/ejecución, normatividad, y evaluación/control.

También se construye un capítulo específico en el informe final para dar respuesta a las preguntas orientadoras del componente 1 (las del componente 2 son respondidas en el PTIR). Aun así, y tal como se puede evidenciar en varios de los fragmentos de la entrevista realizada a MinCIT en la sección de efectividad del proceso de evaluación del presente capítulo, los usuarios previstos de la evaluación tuvieron problemas para comprender las recomendaciones, y su relación lógica con las conclusiones elaboradas. Por esta razón, el estándar de claridad del análisis fue parcialmente satisfecho.

En el informe final se incluye un resumen en su inicio, en el que se realiza una descripción de la metodología y una síntesis de los principales hallazgos (pp. 13-20). Igualmente se construyó una ficha técnica de la evaluación (G/Exponencial, 2016a) que contiene: descripción de la intervención, objetivo general de la evaluación, metodología de la evaluación, principales resultados y conclusiones y las recomendaciones. Nuevamente, la ficha técnica satisfizo mucho mejor los componentes del estándar de claridad y representatividad del resumen, que el resumen ejecutivo.

También se realiza una adecuada descripción del contexto de la intervención. Para esto, se toman en consideración los compromisos adquiridos por Colombia en 1995 al ingresar a la OMC (G/Exponencial, 2016b, p. 22), el número de solicitudes de medidas de defensa comercial que se han realizado, y su comportamiento con el transcurso de los años (p. 26), la incidencia de la reglamentación de estas medidas en los diferentes sectores industriales, la evolución del conocimiento del sector privado de los mecanismos de defensa comercial y su impacto en la presión que se ejerce sobre la capacidad institucional de MinCIT, y la presión de China y la negociación de distintos acuerdos de comercio internacional con economías industrializadas y competitivas (EEUU, Canadá, Unión Europea, Turquía y Corea) que ejercen sobre los sectores productivos internos (p. 27).

Otros elementos importantes con los cuales se realiza la contextualización de la intervención son la descripción de los acuerdos comerciales vigentes (p. 43), la relación entre los compromisos internacionales de Colombia con su ordenamiento jurídico (p. 44), el marco legal a nivel multilateral de la OMC, a nivel de la Comunidad Andina, a nivel nacional sobre Dumping y Salvaguardias y jurisprudencia relevante de la OMC (p. 40). Lo referente a la participación de los actores implicados en la política es abordado a lo largo de la evaluación. Con todo esto, el estándar de contexto de la intervención fue satisfecho.

De todas las evaluaciones que fueron seleccionadas como casos particulares de procesos de evaluación, esta es la que mejor reconstruye la cadena de valor de los procesos de las medidas antidumping y salvaguardia. En total se realizan 12 esquemas de cadena de valor con base en los decretos 152 de 1998, 1407 de 1999, 2550 de 2010, 1750 de 2015. En estos esquemas se incluyen descripciones a profundidad de las entradas, los actores competentes, las actividades, las salidas, y los resultados de los procesos evaluados, y se subdividen en los que grafican los procesos de manera agregada, y los de los subprocesos particulares. Con todo esto, el estándar de lógica de la intervención fue plenamente satisfecho.

En cuanto a la metodología descrita en el informe final, esta contiene el marco conceptual de los elementos que conforman una cadena de valor, una caracterización de la metodología de análisis de procesos denominada “Datleneco”, los instrumentos, variables, indicadores y fuentes de información. La selección de esta metodología se justificó en que esta “… recoge información sobre operaciones en un entorno institucional con una dinámica propia de la complejidad inherente a la administración pública, la cual se caracteriza por relacionar múltiples instituciones en una cadena de valor” (G/Exponencial, 2016b, p. 73). También porque tiene la capacidad de “… observar múltiples decisiones especializadas, tomadas al interior de distintos subsistemas interrelacionados” (p. 73), y de detectar eslabones críticos en la cadena de valor.

El nombre “Datleneco” hace referencia a las áreas de decisión en la administración pública: D de Dirección que se refiere al área de decisión cuyo resultado son los productos finales; A es el área de decisión que define la entrega de recursos y su ubicación; T de tecnología, define el cómo se hacen los productos; L de logística define los recursos para el cumplimiento u obtención del producto final; E de ejecución, define con quién y cuándo se realizan las actividades planificadas; N de normativa, define las reglas que aseguran la convergencia de los comportamientos de las personas; E de evaluación, define los indicadores que miden la materialización del proceso y CO de control, define la verificación del cumplimiento de las condiciones para operar, prever errores, corregir fallas y gestionar lo requerido (pp. 74-75).

Posteriormente, se describen los instrumentos que fueron utilizados, la unidad de investigación a la que se le aplicaron, y la(s) medida(s) que ayudan a caracterizar (p. 77). En total, se realizaron 36 entrevistas a 6 expedientes de procesos estudiados, y dos entrevistas de un expediente adicional que sirvió como prueba piloto del trabajo de campo (p. 82). También se describen los entrevistados por tipo de actor (p. 83). No se realizan menciones acerca de los cambios o límites de la metodología utilizada. La metodología presentada en el informe de evaluación, en síntesis, satisfizo el estándar de explicación de la metodología, y de validez y fiabilidad de las fuentes de información, pero no el de mención de cambios o límites de la metodología.

Además de evidenciarse que se consultó a todo el conjunto de actores relevantes, y que se ven claramente los criterios de su identificación y selección, se observa en el informe final el cumplimiento de los lineamientos formulados en el pliego de condiciones definitivo, relacionados con el anonimato de la información recolectada a lo largo de la consultoría, en orden de proteger los derechos a la intimidad, el habeas data, y buen nombre (DSEPP-DNP, 2015h, p. 5); y con la confidencialidad de los productos (p. 45). Esto significó el cumplimiento del estándar de consulta y protección de los implicados.

No se observan menciones acerca de discrepancias entre miembros del equipo evaluador, sin embargo, en el PTIR se describe que estas discrepancias fueron solucionadas en medio de reuniones internas del equipo evaluador, cuyo caso emblemático fue la discusión alrededor de la pertinencia o no de la realización de ajustes a la normatividad actual (p. 9), ante lo cual, se llegó a la decisión de no recomendar estos ajustes por sus implicaciones en la operación de las medidas. Adicionalmente, se aclara que parte de la metodología utilizada para la construcción de las recomendaciones de la evaluación, consistió en diversas reuniones internas del equipo evaluador a lo largo de todo el proceso de evaluación. Con esto, si bien no se aclararon estas discrepancias, si hubo el espacio para solucionarlas, por lo que el estándar de mención de discrepancias entre miembros del equipo evaluador si fue satisfecho.

Finalmente, en el contrato de la consultoría se estableció como fecha para la entrega del informe final el 19 de octubre de 2016, y para la entrega del PTIR el 30

de noviembre de 2016 (DNP, 2016c, p. 5). En correspondencia, el informe final tiene fecha de octubre de 2016, y el PTIR de noviembre del mismo año. Esto significa que el estándar de evaluación dentro del presupuesto y el tiempo previsto si fue satisfecho en este proceso de evaluación.

9.2.4 Reglas de agregación y control

Al igual que en los demás contratos, las reglas de agregación se concentraron en brindarle al DNP el control suficiente para garantizar la adecuada construcción de los productos. Entre estas reglas se encuentra: entrega de informes técnicos que sean solicitados por el supervisor el contrato –DNP- en un tiempo no mayor a ocho días hábiles (DSEPP-DNP, 2015h, p. 4); expedir el certificado de cumplimiento a satisfacción de los productos, necesario para el pago al consultor; mecanismo de cobertura del riesgo mediante contrato de seguro en póliza, patrimonio autónomo o garantía bancaria, para cubrir los riesgos de: cumplimiento del contrato, calidad del servicio y pago de salarios, en consonancia con la Ley 80 de 1993, Ley 1150 de 2007 y Decreto 1082 de 2015; y multas en caso de mora o incumplimiento por parte del consultor (DNP, 2016c, p. 8).

También se estipuló que el Comité de Seguimiento sería el encargado de realizar comentarios sobre la calidad técnica de los productos, y recomendar parámetros de calidad para su elaboración efectiva (DSEPP-DNP, 2015h, p. 49), y su secretaría técnica (ejercida por la DSEPP) se encargó de recibir los comentarios del Comité Técnico "…consolidarlos, depurarlos y darles el tratamiento respectivo "(p. 50). Al igual que en anteriores evaluaciones, el equipo evaluador contó con una semana para entregar el producto con los ajustes solicitados por el Comité Técnico (p. 4). Estos elementos garantizaron el cumplimiento de los estándares de control de calidad y de inclusión de observaciones de los implicados.

9.2.5 Reglas de posición

Frente a los estándares que entran como reglas de posición, pasa algo similar al resto de las evaluaciones. Varios factores entran en disputa con la independencia del evaluador: la posibilidad de activación de los mecanismos de cobertura de riesgo; la activación de la cláusula correspondiente a las multas por mora o incumplimiento; margen de discrecionalidad restringido, y un tiempo muy limitado para defender sus productos en caso de una discrepancia por parte del evaluador (DSEPP-DNP, 2015h, p. 50); y riesgo de retrasos en el pago al consultor, el cual solo se realiza una vez y se incluyen todos los comentarios del Comité de Seguimiento.

En contraste, se estableció la cláusula correspondiente a la autonomía del consultor, en donde se establece que este actuará en completa autonomía técnica y administrativa (DNP, 2016c, p. 8). Por otra parte, para observar la naturaleza de los comentarios, que daría una idea del grado en que se incidió negativamente la

independencia del evaluador, se puede resaltar lo siguiente de los conceptos técnicos de los productos:

> "… el documento enfatiza mucho en la primera fuente (entrevistas) y se encuentran muy pocas referencias de las demás fuentes a lo largo del documento. Esto le resta contundencia a las conclusiones, pues parecen basadas en elementos de opinión únicamente. Se solicita entonces fortalecer los análisis con el criterio del evaluador a partir de lo encontrado en las revisiones propias de otras fuentes de información" (DSEPP-DNP, 2016d, p. 2)
>
> "El estudio para "identificar las fortalezas y debilidades de las investigaciones", no corresponde a una valoración técnica adecuada por parte del consultor, lo que se refleja en algunas de las conclusiones (…) se fundamenta mucho en la primera fuente (entrevistas) en particular las del sector privado" (p. 2).
>
> "Se evidenció que este producto estaba muy desvinculado del anterior. La consecuencia de esto es que muchas recomendaciones de la lista parecen tener un sustento débil, o no tenerlo" (2016e, p. 2).
>
> "Se recomendó robustecer las recomendaciones y en caso de ser necesario fusionar varias de ellas" (p. 2)

Como se verá más adelante, 5 de 9 de las recomendaciones de esta evaluación tienen evidencia de ser justificaciones de decisiones previas, siendo la evaluación en la que más se evidencia este fenómeno. Aun así, como se puede observar en el tipo de comentarios que fueron realizados por el Comité de Seguimiento, este fenómeno no se relacionó con un escaso grado de independencia, puesto que los comentarios no se dirigen a modificar valoraciones o recomendaciones realizadas, sino a aumentar el rigor de estas. Esto significa que los estándares de independencia del evaluador, y de proceso evaluativo libre y abierto fueron satisfechos en este proceso de evaluación.

9.2.6 Reglas de información

Con relación a los estándares que funcionan como reglas de información, también sucedió algo muy parecido a la mayoría del resto de procesos de evaluación estudiados. El pliego de condiciones definitivo funciona como documento de términos de referencia, en este se presentan todos los componentes requeridos por el estándar de documentación acerca de la finalidad y expectativa, así mismo, en correspondencia con el proceso interno del DNP de diseño de la evaluación, se constató con las entrevistas realizadas que este documento surgió de un consenso con MinCIT.

Por otro lado, en cuanto al informe final de la evaluación, y en correspondencia con el estándar de claridad del análisis, se especificaron lineamientos en el pliego de condiciones definitivo orientados a que el producto fuera comprensible (DSEPP-DNP, 2015h, p. 59), y se observa un pleno cumplimiento de estos en todo el informe final, por lo que el estándar de informe de evaluación fue satisfecho.

Finalmente, en cuanto a la difusión de los productos de la consultoría, se requirió realizar una presentación de los resultados de la evaluación ante los directivos de las entidades que conforman el Comité de Seguimiento (p. 4), en correspondencia, en la entrevista realizada a MinCIT se manifestó: "… hubo una presentación ejecutiva ante las altas instancias, el ministerio, viceministro, ministro, asesores" (Subdirectora de Prácticas Comerciales, MinCIT). Adicionalmente, esta estrategia de difusión incluyó la realización del PTIR, la ficha técnica mencionada, presentación extendida, presentación simple e infografía (DSEPP-DNP, 2015h, p. 61). En el portal de SINERGIA se puede observar que se hizo difusión externa del informe final, la ficha técnica, el resumen ejecutivo, y la infografía. Esto satisfizo el estándar de difusión de la evaluación.

9.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping

En cuanto a los patrones de interacción y la coordinación que se dio entre los actores del proceso de evaluación, en primer lugar, hay que destacar la ausencia de la mayoría de los miembros del Consejo Superior de Comercio Exterior en el Comité de Diseño y en el Comité Técnico de la evaluación, de 11 miembros que tiene este consejo, únicamente se involucró como usuario primario a MinCIT. Sin embargo, se tuvo cuidado de que las recomendaciones estuvieran dirigidas a MinCIT (únicamente una tiene como responsable a actores que no fueron involucrados como usuarios de la evaluación). Por lo que hubo un buen nivel de hallazgo e involucramiento de los usuarios que se consideraron necesarios para mejorar las medidas

Respecto al equipo de trabajo conjunto entre usuarios y evaluadores, y a su dinámica de trabajo, se tiene información dividida en cuanto a la reactividad del evaluador y del DNP frente a las preocupaciones y demandas del usuario previsto. Evidencia que muestra fuertes límites en este sentido, resalta falencias en la verificación del estado de la implementación de las medidas de defensa comercial y problemas en la claridad de las recomendaciones:

> "Mira, algo que sí te puedo decir es que ellos hicieron muchas de esas recomendaciones sin verificar exactamente lo que nosotros teníamos. El

> estudio fue más dirigido a hacer la encuesta, al sector privado, a las empresas, a los miembros del comité de prácticas, a consultores externos, pero que ellos se dedicaran a revisar puntualmente o verificar lo que nosotros teníamos y lo que nosotros hacíamos, no" (Subdirectora de Prácticas Comerciales, MinCIT).

En el mismo sentido, se observa que en el PTIR se establece como hallazgo de la evaluación: "Falta un reglamento interno para la realización de seguimientos" (p. 32), ante lo que se recomienda crear un reglamento de seguimiento, sin embargo, en MinCIT aclaran:

> "… no entendimos, nosotros tenemos los exámenes, tenemos las revisiones, porque lo tenemos en el plan operativo nuestro, de hacer monitoreo a los productos que tienen medida. Entonces, no fue claro y el ministerio realmente no tomó esa recomendación como esencial para nosotros" (Subdirectora de Prácticas Comerciales, MinCIT).

Estos límites en la reactividad del evaluador frente a las preocupaciones y demandas de los usuarios reflejan que el interés de estos últimos no fue prioritario en todo momento a lo largo del proceso de la evaluación. También que a pesar de llegar en estos puntos a conclusiones erróneas y a recomendaciones irrelevantes (e incluso carentes de claridad), el Comité de Seguimiento no sirvió a la corrección de estas fallas particulares a pesar de tener el marco institucional para hacerlo. Sin embargo, también existe evidencia que le hace contrapeso a lo anterior. Por una parte, se aclaró que la evaluación surgió de un interés por parte de MinCIT:

> "… aplicar medidas antidumping requiere de un estudio, de un análisis y un desarrollo. Entonces el sector privado venía cuestionando que se requería mucha información, que nos demorábamos mucho, que no se aplicaban las medidas antidumping (…) entonces la viceministra [de Comercio Exterior] que estaba de turno en ese momento solicitó la ayuda del DNP para que se hiciera una evaluación de estos procedimientos a ver qué, dónde estaban los cuellos de botella (…) o espacios donde uno pudiera mejorar el proceso" (Subdirectora de Prácticas Comerciales, MinCIT).

Como se pudo observar con los estándares de fundamentación y finalidad de la evaluación, objeto y alcance de la evaluación, y objetivos específicos, la evaluación efectivamente se direcciona hacia este uso. También esto se ve reforzado por la naturaleza de las recomendaciones realizadas, tal como se verá en la siguiente sección. Adicionalmente, y en un sentido general, en MinCIT manifiestan que se satisfizo su demanda general de información al ser indagados al respecto:

> "… digamos, la revisión frente a otros países era importante para saber si lo estábamos haciendo muy diferente, muy mal o, o en qué se podía mejorar. Y yo creo que hay una cosa buena y es en el sentido de que no estamos haciendo las cosas tan mal o tan diferente a como se viene haciendo en el mundo. Sí, hay procesos, hay cosas que hay que mejorar: el fortalecimiento

> del equipo de trabajo es algo fundamental, porque la carga de trabajo que fue una cosa que me gustó que hizo el estudio, fue hacer una medición de carga y las conclusiones del fortalecimiento del equipo surgieron de esa medición de la carga de trabajo (…) hay muchas cosas que nosotros también tenemos a cargo y no solamente la investigación. Entonces se veía que había unas cargas de trabajo que también, por ese motivo, había que fortalecer el equipo" (Subdirectora de Prácticas Comerciales, MinCIT).

Como se observa, muchas de las preocupaciones de los usuarios si fueron finalmente abordadas por el equipo evaluador, a pesar de que no fue en la totalidad de los casos. En la misma dirección, al ser indagados acerca de la coordinación que hubo en el Comité de Diseño y en el Comité de Seguimiento se manifestó una percepción positiva, y se encuentran elementos que evidencian reactividad, particularmente en cuanto a temas metodológicos:

> "… La coordinación estuvo bien, estuvo bien porque ellos estuvieron muy atentos, y las personas, sobre todo los técnicos [del DNP] que trabajaron conmigo en ese comité, son personas que ya han adquirido una experticia en estos temas porque participan y son los asesores que vienen a estos comités de triple A como llamo yo el de asuntos aduaneros, arancelarios y el de prácticas comerciales. Y ellos ayudaron mucho en hacer claridad en muchas cosas (…) el diseño para hacer ese estudio se ceñía por unas metodologías, de trabajo, de encuestas, y muchas veces las conclusiones solamente atendían esas respuestas de las encuestas. Entones, eso era una de las discusiones que dábamos en estos comités de coordinación, que no se basaran únicamente en las respuestas. Porque, por ejemplo, el empresario, si tú vas y le preguntas, tiene un interés, y sobre todo los empresarios que ya han participado en estos procedimientos, entonces respondían de acuerdo a sus intereses, pues eso es natural, pero muchas veces hay un sesgo ahí, entonces hay que visionar la respuesta de la encuesta en un contexto más amplio" (Subdirectora de Prácticas Comerciales, MinCIT).

Frente al involucramiento de los usuarios primarios en la interpretación de los hallazgos, también se halla evidencia dividida. Por una parte, en MinCIT se rechazó la idea de que fueran comparados y que se hicieran recomendaciones, con base en buenas prácticas que tomaron por ejemplo paradigmático a la Unión Europea: "…nosotros cuestionamos eso porque ¿por qué Unión Europea que institucionalmente es totalmente diferente a Colombia? Unión Europea es una unión aduanera perfecta, 27 países, que tienen un reglamento, y una comisión europea que es la que define todos estos asuntos" (Subdirectora de Prácticas Comerciales, MinCIT). En efecto, MinCIT solicitó que fuesen comparados con otros países como México, Brasil o Argentina, sin embargo, en el capítulo del informe final de "Referencias internacionales de buenas prácticas" (G/Exponencial, 2016b, p. 90), se evidencia que únicamente se comparó con la Unión Europea, y que, en el PTIR, dos de las recomendaciones se sustentan (por lo menos parcialmente) en esta

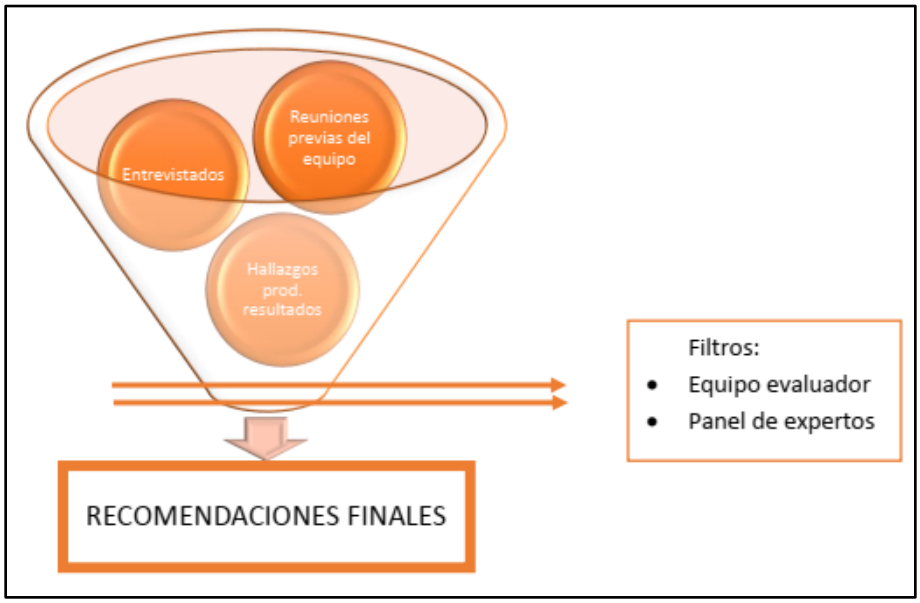
comparación, a pesar de estos contrastes resaltados por MinCIT. A pesar de esto, con respecto a la revisión de los productos se manifestó:

> “Sí, claro, y funcionó durante el estudio, funcionó, y yo participaba, hacía las observaciones frente al estudio que nosotros considerábamos. Cómo funcionaba: nos mandaban los documentos con anticipación para nosotros hacer la revisión y qué observaciones teníamos frente a esos entregables parciales que se hacían. Y nosotros éramos los encargados de hacer observaciones” (Subdirectora de Prácticas Comerciales, MinCIT).

Entonces, se tiene que se realizaron comentarios a los productos de la consultoría, que MinCIT tuvo distintos escenarios en los cuales pudo manifestar sus preocupaciones y demandas, sin embargo, por otro lado, se evidencia que en varias ocasiones estos comentarios no fueron tomados en cuenta por parte del evaluador consultor, a pesar de su grado de pertinencia. Esto conlleva, además, a que los usuarios primarios no se apropien de (por lo menos varios de) los resultados de la evaluación por considerarlos erróneos.

Similar a esto, y con relación al involucramiento de los usuarios en la construcción de recomendaciones, al analizar su metodología de elaboración (G/Exponencial, 2016c, pp. 6-9) se puede constatar que, a diferencia de varios de los anteriores casos, en este caso no se incluyeron los comentarios de los usuarios previstos como una etapa particular para la construcción de las recomendaciones, de modo que su involucramiento en esta etapa fue limitado.

De manera puntual, la misma firma consultora que en otra evaluación había aclarado la incidencia del Comité de Seguimiento en la elaboración de las recomendaciones, especifica las fuentes de las observaciones que fueron tomadas en cuenta para modificar su propuesta de las mismas, dentro de estas se incluyen: recomendaciones suministradas por los entrevistados en el trabajo de campo; recomendaciones obtenidas de diversas reuniones internas del equipo evaluador; recomendaciones obtenidas de las conclusiones del informe de resultados de la evaluación; y el panel de expertos. Incluso se esquematiza esta metodología, en la cual se evidencia la ausencia de los usuarios previstos:



Fuente: G/Exponencial, 2016c, p. 9

Por último, no se encontraron indicadores para hacer seguimiento a la implementación de las recomendaciones, ni se encontró ninguna evidencia acerca de un plan conjunto de seguimiento para este punto. Tampoco acerca de búsqueda proactiva del uso de los resultados de la evaluación por parte del equipo evaluador, al igual que en las demás evaluaciones, por la terminación de la relación contractual con el DNP.

9.4 Análisis de la efectividad de la evaluación del proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping

La evaluación produjo un total de 9 recomendaciones, las cuales se dirigen hacia los procesos e instrumentos, la normatividad y las instancias de recomendación para la toma de decisión de la adopción de las medidas de defensa comercial. La primera recomendación fue mejorar los procedimientos relativos a la solicitud de las medidas de defensa comercial a partir de (G/Exponencial, 2016c, p. 11):

- “Implementar un equipo de atención al usuario y pedagogía”.
- “Ajustar la información publicada sobre las medidas de defensa comercial”.
- “Reglamentar una expresión inicial de interés”.
- “Proporcionar mayor flexibilidad en la entrega de información”.
- “Comunicar la solicitud a los gremios correspondientes”.
- “Generar un reglamento para la etapa de solicitud, y que se haga entrega de documentación únicamente a través de medio electrónico”.

Al ser indagados por su estado de implementación, en MinCIT contestaron:

> “Sobre eso, hay algunas que, parcialmente, eso ya venía funcionando. De hecho, la recomendación se hizo pero ya existía el mecanismo de socialización de una guía que aparece en nuestra página web, en la página web del Ministerio de Comercio. (…) Ahí tú puedes desplegar (…) desde guías, manuales, formatos, que facilitan, que buscan precisamente facilitar la presentación de la solicitud. De hecho, este es uno de los temas que precisamente, en el cual tiene que existir un presupuesto disponible por el cual hay que crear un equipo especial (…) Esto fue una recomendación, de hecho, que surgió para que incluso mejoráramos el lenguaje que creo que más adelante son algunas de esas recomendaciones, el lenguaje de las guías para que lo hiciéramos más amigable, para que la gente pudiera entenderlo más fácilmente, eso se hizo. De esa recomendación, lo que no se hizo fue crear un equipo de atención diferente a lo que es hoy en día la autoridad investigadora (…) Pero el resto pues sí, hay y lo puedes verificar en la página web, aquí te voy a mostrar, uno entra a la página del Ministerio y eso no surgió a raíz de esta recomendación, ya existía. Lo mismo que el tema del aplicativo [para la solicitud], ya se venía trabajando en los ajustes de un aplicativo que en un principio no fue tan amigable, hubo que hacerle muchos ajustes, y el tema también está socializado en la página de internet (…) son herramientas que existen digamos, dadas por los mismos acuerdos pactados por Colombia al ingresar a la OMC, el acuerdo antidumping son acuerdos multilaterales, son reglas y disciplinas que se pactaron en OMC, que le permite a los países, además de la liberación de comercio, aplicar medidas de defensa comercial” (Subdirectora de Prácticas Comerciales, MinCIT).

Se tiene entonces que, de los 6 componentes de esta recomendación, uno no se implementó, y el resto corresponde a compromisos que se tuvieron con la OMC y, por lo tanto, eran actividades que ya estaban en marcha antes de tener los resultados de la evaluación. Esto significa que esta recomendación no generó un cambio en el modelo mental de los usuarios, funciona a modo de justificación de decisiones previas, no mejora el proceso de toma de decisiones, ni refleja efectividad del proceso de evaluación.

La segunda recomendación fue destrabar el cuello de botella de la etapa de investigación a partir de la organización de las audiencias, precisión de fechas de la publicación de documentos, dar la posibilidad a la Subdirección de Prácticas Comerciales de contratar consultorías, construir y publicar manuales, protocolos y guías, y brindar mejoras tecnológicas (G/Exponencial, 2016c, p. 18). En el anterior fragmento ya se había aclarado que MinCIT ya había avanzado en la elaboración de manuales, protocolos y guías. Con respecto a esta recomendación se comentó algo muy similar a la anterior:

> “No, eso está hecho, y lo que pasa es que de esas se buscan los mejoramientos a esos procesos, o a esas, digamos, etapas. Nosotros tenemos la norma, tanto el acuerdo multilateral de la OMC como la norma

> nacional que regula estos procedimientos, permite que haya unas audiencias públicas. Esas audiencias públicas son entre las partes intervinientes. Colombia, por ejemplo, es uno de los países que elaboró unas reglas para la participación en esa audiencia y cómo se desarrolla la audiencia, como para darle un orden. Entonces, cuando una parte dice: 'solicito se convoque una audiencia pública entre partes intervinientes' nosotros miramos si cumple con toda la legalidad, es decir, si es una parte, si ha intervenido en el proceso, si cumple con esas condiciones, o de oficio también lo podemos hacer: se convoca" (Subdirectora de Prácticas Comerciales, MinCIT).

Tal como es caracterizado por la funcionaria, todo lo que se recomienda son actividades que ya se estaban ejecutando, y que se iban modificando con la misma dinámica de la entidad. En este orden de ideas, al estar dirigida la recomendación a mejorar algo que ya estaba siendo mejorado, no se amplía el rango de acciones ni de resultados posibles, por lo que no se mejora la toma de la decisión. La recomendación también funciona para justificar decisiones previas, ya sea de la OMC o de MinCIT, y no refleja efectividad del proceso evaluativo a pesar de que estas acciones se puedan ver en una primera mirada como implementadas.

En tercer lugar, se sugirió mejorar el procedimiento de recomendación y decisión a partir de un tablero de indicadores para las instancias de recomendación, una estrategia de transparencia (rendición de cuentas), y que hubiese concordancia entre el informe técnico que generan las investigaciones y la decisión final (G/Exponencial, 2016c, p. 29). Esta última parte ya está incluida en la normativa y es la justificación y utilidad de la realización de las investigaciones que hace MinCIT, por lo que básicamente se está sugiriendo que se utilicen las investigaciones para lo que fueron hechas. Con respecto a lo demás se comentó:

> "Esa no. Generar un tablero de indicadores (…) definir una estrategia de rendición de cuentas públicas, bueno, sí y no, directamente nosotros tenemos todos estos tableros de indicadores y la estrategia de rendición de cuentas públicas que maneja la oficina que está dentro del Plan Estratégico del ministerio, nosotros estamos ubicados dentro del área del viceministerio, y dentro de eso nos ceñimos a todos los procedimientos. Cada uno de estos mecanismos, y cada una de las actividades, tiene un procedimiento bajo el Sistema Integrado de Calidad que se rige por estos, por este mapa de indicadores, la estrategia de rendición de cuentas pública que es la misma de todo, está integrada en todo el ministerio, a través puede ser de audiencia públicas, de capacitaciones. Muchas cosas de esas, de la rendición de cuentas pública es la política que sigue el ministerio (…) estas fueron las recomendaciones que el ministerio no entendió mucho y que realmente no sabíamos, digamos, cómo enmarcarlas solamente en esta área, porque eso ya está. Está definido como una estrategia del ministerio de la rendición de cuentas públicas y está dentro del mapa de indicadores que tenemos. Nosotros, el procedimiento de dumping está dentro del sistema de gestión de

> calidad del ministerio, de nosotros. Y tenemos que cumplirlo. Y si los indicadores y aplicar los indicadores y todo" (Subdirectora de Prácticas Comerciales, MinCIT).

Con esta recomendación pasó exactamente lo mismo que en las anteriores, se sugirieron acciones que ya existen, por lo que justifican decisiones previas y no mejoran el proceso de toma de decisiones. Aún más problemático, y en correspondencia con lo visto en el análisis de patrones de interacción, es el hecho de que se construyeran recomendaciones sin "… verificar exactamente" (Subdirectora de Prácticas Comerciales, MinCIT) lo que tenía el MinCIT en esos aspectos particulares.

La cuarta recomendación consistió en instaurar ajustes al procedimiento de seguimiento a las medidas de defensa comercial, puntualmente en temas de mejora en los mecanismos de socialización de los resultados, estudio de la situación de control aduanero de bienes con medidas de defensa comercial, que el desdoblamiento arancelario se autorice con anterioridad a la resolución de imposición de medidas, informes anuales de seguimiento, monitoreo del cumplimiento de los planes de ajuste, publicación de expiración de las medidas (G/Exponencial, 2016c, p. 31). Frente a su implementación se contestó:

> "En nuestro país, nosotros nos ceñimos, como todo el mundo, al Sistema Armonizado de la Organización Mundial de Aduanas, entonces ellos emiten unas enmiendas cada dos años o cada cuatro años, en donde se hace la clasificación arancelaria de un producto y la subpartida que le corresponde, [es] un código que las identifica. En muchas ocasiones, por una misma subpartida, pueden ingresar varios productos, que no son los mismos, entonces ¿qué pasa? Esto no lo acogió el ministerio como tal porque es una mezcla de algo que no depende directamente del ministerio, porque el tema de los desdoblamientos se define ante el comité triple A, o sea un Comité de Asuntos Aduaneros, Arancelarios y de Comercio Exterior, no dentro de la investigación antidumping como tal. Lo ideal es que en una investigación antidumping en donde se cuestione que por la misma subpartida vienen muchos productos diversos, existiera primero un desdoblamiento, esa es lo ideal. Pero no siempre favorece al sistema de control de aduanas, porque la aduana de nuestro país, desafortunadamente, no tiene la capacidad suficiente, muchas veces, ni los laboratorios suficientes, para poder distinguir un producto del otro" (Subdirectora de Prácticas Comerciales, MinCIT).

En este sentido, al realizar esta recomendación no se verificó que los usuarios previstos no tienen la competencia para implementar esta recomendación. Los miembros del Comité Triple A faltantes son: Ministerio de Hacienda y Crédito Público, Ministerio de Agricultura y Desarrollo Rural, Ministerio de Minas y Energía, DIAN, Superintendencia de Industria y Comercio, y los delegados del Consejo Superior de Comercio Exterior. Adicionalmente, se observa que no se tuvo en cuenta las limitaciones en la capacidad institucional. Con relación al resto de los elementos de la recomendación se manifestó:

> "… es muy difícil el tema de la socialización de los resultados. Lo hacemos, ya te estoy mostrando la página, te estoy mostrando no solamente eso, sino a nivel de capacitación [a los beneficiarios] siempre estamos hablando de las medidas. Los sistemas de control de aduanas no dependen directamente del ministerio, ni de nosotros como área acá. Entonces, nosotros veíamos que no había como algo sólido en esas recomendaciones de por qué y cuál fue la razón para hacer este tipo de recomendación (…) Digamos que ahí hay mucha mezcla de cosas. Entonces, el ministerio dice: todas las resoluciones de aplicación de medidas donde se les pone una vigencia, y a partir del momento en que se cuenta la vigencia de las medidas, se publican en el diario oficial, esa es una norma de carácter general y por ley nosotros tenemos que hacerlo. No es que el estudio [la evaluación] me lo recomiende, es que es por ley, obligatorio publicar las resoluciones de carácter general. Y así fueran de carácter particular, aquí hay una publicidad que en la página web también del ministerio (…) Entonces, yo no entendí, una de las cosas que se le cuestionaron al estudio fue que, precisamente, no hay una causa de esa recomendación, no entendemos de dónde surge y por qué surge" (Subdirectora de Prácticas Comerciales, MinCIT).

De esto se puede inferir que además de desconocer la normatividad en la cual ya está integrada la recomendación, debido a la combinación de múltiples elementos disímiles, no hubo suficiente claridad en la relación lógica entre conclusiones y recomendaciones para los usuarios previstos, a pesar de que se hace el esfuerzo en esta dirección dentro del PTIR, tal como se evidenció en el estándar de claridad del análisis. El último componente de la recomendación fue elaborar un reglamento particular para el seguimiento (G/Exponencial, 2016c, p. 31), frente a esto se comentó:

> "… decíamos ¿pero el seguimiento en qué sentido? (…) esas medidas las aplica el Ministerio de Comercio y las hace efectivas la DIAN. La DIAN tiene los mecanismos informáticos para introducir a las importaciones que lleguen de ese determinado producto y hacerle efectiva esa medida. Entonces, la DIAN tiene unos mecanismos de seguimiento para la aplicación de esa medida y eso (…) Las medidas se imponen por cinco años. Cuatro meses antes del vencimiento de esa medida (…) el productor nacional viene acá y dice 'mire, yo necesito que me prolongue esa medida'. Cuando nosotros hacemos ese examen, nosotros miramos cómo fue el escenario en que estuvo aplicada la medida, que fueron casi cinco años, y cómo sería un daño prospectivo, porque aquí se mira hacia futuro (…) Entonces, no entendimos, nosotros tenemos los exámenes, tenemos las revisiones, porque lo tenemos en el plan operativo nuestro, de hacer monitoreo a los productos que tienen medida, entonces no fue claro y el ministerio realmente no tomó esa recomendación como esencial para nosotros" (Subdirectora de Prácticas Comerciales, MinCIT).

Igual que antes, esta recomendación no verifica lo que ya se estaba haciendo en los componentes particulares de los procedimientos de la intervención a los que buscaba incidir. Claramente, este componente de la recomendación al ser irrelevante no fue implementada. En síntesis, los componentes de la recomendación que no se implementan por irrelevancia, justifican decisiones ya tomadas en el pasado, por lo que, en ninguno de los dos casos, se revela efectividad del proceso evaluativo.

En quinto lugar, y relacionado con el tema del seguimiento a las medidas de defensa comercial, se sugirió incluir al interior del reglamento interno del seguimiento que dentro de los compromisos trazados por el solicitante, cuando se le otorga una medida de salvaguardia, la Subdirección de Prácticas Comerciales tuviera derecho a "…acceder a toda la información que considere necesaria para realizar sus investigaciones, de manera permanente, mientras esté vigente la medida" (G/Exponencial, 2016c, p. 43). Acerca de esto se respondió:

> "… eso sí fue algo también que logramos a partir del 2015, 2016, que también puede estar en las recomendaciones, pero ya lo habíamos iniciado. Digitalizar los expedientes públicos (…) se levanta un expediente público y confidencial. El público se digitaliza y se mantiene en la página web, le voy a mostrar éste, por ejemplo, de papa congelada. Papa que es contra unos países europeos. Y acá abajo, está todo el expediente, lo puede revisar desde acá. Ya no tenemos, como teníamos antes, que la gente venía aquí a sacar fotocopias del expediente físico y todas esas cosas, no, ya lo pueden consultar directamente en la página web. Pero toda la información y la información de la investigación, desde su apertura, como todo el proceso, se publica, se va publicando, y ahí está. Nosotros enviamos cuestionarios a los exportadores y entonces se cuelgan aquí, estos son cuadros estadísticos, porque esta es una investigación de un contexto económico pero también jurídico, el proceso legal" (Subdirectora de Prácticas Comerciales, MinCIT).

Como se observa, todos los componentes de la recomendación ya se habían comenzado a implementar antes de tener los resultados de la evaluación. En este sentido, esta recomendación no es el modelo mental de los usuarios, justifica decisiones ya tomadas por lo que no mejora el proceso de decisiones, y no refleja efectividad del proceso de evaluación.

La sexta recomendación fue establecer mejores lineamientos para el manejo de la confidencialidad de la información, que estén direccionados a establecer las mismas condiciones para todos los beneficiarios en cuanto a la naturaleza de la información que pueda ser confidencial o no (G/Exponencial, 2016c, p. 36). Como se observó, los lineamientos para la recolección y almacenamiento de la información de los beneficiarios para la realización de la investigación por parte de MinCIT, se implementaron al momento de poner en funcionamiento el expediente que se describió como "confidencial".

Haciendo un examen más profundo se encontró que el estatus de confidencialidad está reglamentado en el artículo 41 del Decreto 1750 de 2015, y en el documento: *Dumping y Salvaguardias: Guía de uso del aplicativo Web – Peticionarios*, realizado por la Subdirección de Prácticas Comerciales[64] (p. 6), en estos documentos, se aclara que debe ser el peticionario quien decida y aclare la información que debe ser tratada como confidencial. Este artículo sigue vigente hasta la fecha, y la guía fue recolectada en el portal de MinCIT en mayo de 2019, por lo que esta recomendación no fue implementada.

En séptimo lugar se recomendó elevar el rango de la Subdirección de Prácticas Comerciales a dirección dentro de MinCIT (G/Exponencial, 2016c, p. 38). Las razones por las cuales se recomendó esto fueron que la subdirección debe lidiar con otros actores de alto perfil dentro del gobierno, y a que se encontró escasez de recursos de todo tipo para hacer las investigaciones correspondientes. Sobre esto, se contestó lo siguiente:

> "Eso sí, el ministerio lo pasó pero le respondió al DNP con demandas de dinero, porque eso requiere es presupuesto. Lo de fortalecer el equipo investigador sí se hizo, de hecho, reforzaron el equipo, antes éramos muy pocos, y hoy seguimos siendo pocos pero, se reforzaron con cinco personas más. No elevarnos a dirección, pero sí fortalecer el equipo. De hecho, ahorita están mirando cómo se fortalece nuevamente el equipo" (Subdirectora de Prácticas Comerciales, MinCIT).

En efecto, se encontró una solución al problema caracterizado por la evaluación, sin embargo, la recomendación producto del proceso de evaluación no fue implementada como tal.

La octava recomendación consistió en que la investigación realizada por MinCIT debería ser suficiente para tomar la decisión de adopción o no de una medida de defensa comercial, por lo que se sugirió limitar los proceso que son trasmitidos a instancias de recomendación y decisión, donde solo se discutan casos que afecten el interés nacional (G/Exponencial, 2016c, p. 40). Al respectó se manifestó:

> "… si ellos o que dicen es que no debería la instancia, la autoridad investigadora, debería ser la misma instancia decisoria. Solamente, digamos que el ministerio, por ejemplo, con esta área, la dirección o el viceministro, tomarán la decisión y sólo en el caso de que existiera un interés general muy especial se llevara a una instancia de recomendación (…) Hoy en día funciona que nosotros aquí, este grupo, trabaja digamos que somos muy

[64] Este documento se encontró en el link: http://www.mincit.gov.co/getattachment/mincomercioexterior/defensa-comercial/guia-de-uso-del-aplicativo-peticionarios/ guia-de-uso-del-aplicativo-peticionarios.pdf.aspx

> dados a desarrollar toda la parte técnica de la investigación (…) En la etapa preliminar, se pueden adoptar derechos provisionales de manera previsoria y lo adopta el director de comercio, con base en el informe técnico nuestro. Y la decisión final sí va a un Comité de Prácticas Comerciales que evalúa el informe técnico y hace una recomendación. En este momento, el mismo director está encargado de adoptar esa decisión final, que no se puede apartar de la recomendación del comité" (Subdirectora de Prácticas Comerciales, MinCIT).

Entonces, se tiene que el procedimiento sigue funcionando de igual manera a lo que se realizaba antes de tener los resultados de la evaluación, no se limitaron los procesos a instancias de recomendación, y el informe técnico no es en este momento el único insumo que se toma en cuenta para tomar la decisión alrededor de las medidas. Si todo esto sigue igual, la recomendación no fue implementada.

Por último, para mejorar la dinámica de la instancia de recomendación y decisión, se sugirió la apertura de espacio a las partes involucradas en instancias de recomendación y decisión (G/Exponencial, 2016c, p. 44); suprimir las funciones del Consejo Superior de Comercio Exterior, como instancia de recomendación y toma de decisión en los procesos de defensa comercial, a razón de que no se considera necesaria su participación (p. 45); y la elaboración del reglamento interno de las instancias de recomendación y decisión. Al respecto se puede encontrar que el artículo 27 del Decreto 2553 de 1999 sigue vigente, el cual aclara que dentro de las funciones del Consejo esta: "Examinar y recomendar al Gobierno Nacional la adopción de normas generales sobre prácticas desleales y restrictivas de comercio internacional y salvaguardia". Sin embargo, acerca del denominado Comité de Prácticas Comercial, se comenta:

> "… el comité es interinstitucional, se conforma por los dos viceministros, que es el de comercio exterior y el de desarrollo empresarial, que casi siempre es industria, pero cuando no es industria sino que es tema, por ejemplo, papa, participa es el viceministro de agricultura. O si es un tema de minas pues participaría el viceministro de minas, depende del producto participa el viceministro delegado. Y participa la directora sectorial de planeación, participa la directora de gestión de aduanas, participa el superintendente delegado de la SIC, dos asesores del consejo superior. Esos son los miembros del comité de prácticas (…) ellos evalúan nuestro estudio técnico y hacen la recomendación. Muchas veces, ellos piden mayor información entonces ahí se trabaja esa información adicional" (Subdirectora de Prácticas Comerciales, MinCIT).

Así mismo, en otro punto de la entrevista se aclaró: " [tenemos] el reglamento interno del comité que también resultó de una de esas recomendaciones". En este orden de ideas, el primer y el tercer componente de la recomendación si fueron implementados, a pesar de que no se ha hecho un cambio normativo frente al decreto que establece las funciones del Consejo Superior de Comercio Exterior, por esta razón, puede considerarse que si se mejoró la dinámica de la instancia de

recomendación y decisión en el sentido en el que fue sugerido por la recomendación de la evaluación.

9.5 Conclusiones

En síntesis, únicamente existieron problemas en cinco de los estándares de calidad: participación de los actores relevantes, consideración de evaluación conjunta, enfoque asociativo, claridad del análisis, mención de cambios o límites de la metodología. Esto significa que, en términos de estándares, los mayores problemas se encuentran en las reglas de frontera, al presentarse ausencias importantes de actores involucrados en la intervención. Esto incidió negativamente en la posibilidad de implementar recomendaciones, particularmente de las que implicaban cambios normativos a nivel de decretos.

Con respecto a los patrones de interacción con MinCIT, como usuario previsto, se cumplió con lo necesario en cuanto a hallazgo e involucramiento de usuarios previstos. Se encontró evidencia dividida en cuanto a la reactividad del equipo evaluador y del DNP frente a las preocupaciones y demandas de los usuarios, puntualmente se puede decir que en la mayoría de los casos si hubo reactividad, pero con excepciones que no siempre fueron bien justificadas.

Esta tendencia dividida es igual frente al involucramiento de los usuarios en la interpretación de los hallazgos. En contraste, se encontró evidencia negativa en cuanto al involucramiento de los usuarios previstos en la construcción de las recomendaciones. No se encontró evidencia acerca de un plan conjunto de seguimiento a la implementación de los hallazgos de la evaluación, ni acerca de una búsqueda proactiva del uso por parte del equipo evaluador.

De nueve recomendaciones que produjo el proceso de evaluación, cinco funcionan a modo de justificación de decisiones previas, por lo que no mejoran el proceso de toma de decisiones al no tener ninguna incidencia en el cambio de las acciones y los resultados posibles. Tres recomendaciones fueron descartadas por lo que no fueron implementadas. Únicamente una fue implementada. Cabe mencionar que varias de estas recomendaciones resultaron ser irrelevantes por desconocimientos acerca de lo que ya se estaba implementando, de la dinámica de la intervención o de la normatividad.

Finalmente, el último producto de la evaluación estuvo disponible el 28 de diciembre de 2016 (DSEPP-DNP, 2016e), lo cual fue seguido de cambios institucionales como la Resolución 1128 de junio de 2017 de MinCIT, en la cual se establece el reglamento interno del Comité de Prácticas Comerciales, nuevas guías metodológicas y procedimentales internas para los procesos de investigación de las medidas de defensa comercial, la implementación del aplicativo informático para el

trámite electrónico, y el aumento de cinco personas para el grupo de trabajo de la Subdirección de Prácticas Comerciales. Estas modificaciones reflejan que la evaluación se realizó en un momento de rediseño de las medidas, y con un adecuado grado de apertura de los funcionarios a cargo de la política frente a cambios institucionales en la misma.

10. Efectividad de la Evaluación de Resultados de la Política de Prevención y Control de la Contaminación del Aire

10.1 Presentación de la Política de Prevención y Control de la Contaminación del Aire (PPCCA)

La PPCCA actualmente cuenta con dos documentos principales de diseño de política pública: El CONPES 3943 de 2018, y el documento de formulación de política pública construido por el que en su momento era el Ministerio de Ambiente, Vivienda y Desarrollo Territorial en 2010. Teniendo presente que la evaluación de la política pública realizada en el marco de SINERGIA fue llevada a cabo entre los años 2016 y 2017, en esta presentación se muestra únicamente el diseño del documento de política del 2010.

La política tiene por objetivo general "Impulsar la gestión de la calidad del aire en el corto, mediano y largo plazo, con el fin de alcanzar los niveles de calidad del aire adecuados para proteger la salud y el bienestar humano, en el marco del desarrollo sostenible" (Min. Ambiente, Vivienda y Desarrollo Territorial, 2010). Esto se buscó obtener a partir de los siguientes objetivos específicos (p. 31):

- "Regular los contaminantes de la atmósfera que pueden afectar la salud humana y el bienestar de la población, fijando niveles adecuados para proteger la salud de la población y el bienestar humano".
- "Identificar las principales fuentes de emisión de los contaminantes que afectan la salud humana y el bienestar de la población".
- "Establecer, promover y fortalecer las estrategias para prevenir y minimizar la generación de emisiones de contaminantes y de ruido a la atmósfera".
- "Fortalecer espacios de coordinación, participación y capacitación que involucren a los diferentes actores relacionados con la prevención y control de la contaminación del aire".
- "Continuar la implementación de compromisos internacionales adquiridos por el país e incrementar el aprovechamiento de las oportunidades que ofrecen

los acuerdos multilaterales sobre medio ambiente, relacionadas con prevención y control de la contaminación atmosférica".

En la parte programática de dicho documento de política, se construye una matriz a modo de "plan de acción", la cual desglosa cada uno de los objetivos específicos de la PPCCA por: objetivo, meta, indicador, actividad, plazo, y responsable de implementación. El primer objetivo específico se desglosa en dos metas y quince actividades; el segundo se desglosa en tres metas y diez actividades; el tercero en siete metas y 43 actividades; el cuarto en 5 metas con 12 actividades; y el quinto en 5 metas con 16 actividades.

En este documento se establece que aparte del antiguo Ministerio de Ambiente, Vivienda y Desarrollo Territorial (cuya responsabilidad fue trasladada al Ministerio de Ambiente y Desarrollo Sostenible –MADS), dentro de las entidades que se deben encargar de la implementación de estas actividades se encuentra: Ministerio de Salud y Protección Social, IDEAM, entes territoriales, Ministerio de Transporte, Ministerio de Hacienda, MinCIT, Ministerio de Minas y Energía, DNP, entre otros, indirectos que son aclarados en la matriz como los gremios y la academia. La mayoría de estos actores restantes no fueron incluidos dentro del proceso de formulación de la política en 2010.

10.2 Caracterización de las situaciones de acción del proceso de evaluación de la Política de Prevención y Control de la Contaminación del Aire

10.2.1 Reglas de alcance y resultados potenciales

Este proceso consistió en una evaluación de resultados de la PPCCA para determinar la efectividad de las medidas de prevención y control, y para elaborar recomendaciones que permitieran mejorar la gestión de la calidad del aire en el país (DSEPP-DNP, 2016f, p. 1). Dentro de los elementos que se subrayan para justificar la pertinencia de la evaluación, están los costos por muertes prematuras y en enfermedades, que ascendieron del 2002 al 2010 de 4 billones a 10 billones de esos (p. 2); la necesidad de analizar la efectividad de la PPCCA ante la persistencia de concentraciones de contaminantes en la atmósfera por encima de los estándares establecidos después de 6 años de implementación; la necesidad de evaluar la política a mediano plazo, establecida en el documento de política pública; y por último, la necesidad que representa la realización de una evaluación de la efectividad de la PPCCA (p. 2), para tomar mejores decisiones.

El alcance de la evaluación fue determinado a partir de tres componentes. El primero se dirigió a la caracterización de las zonas o regiones objeto de estudio a partir de variables relacionadas con: condiciones ambientales, económicas, sociales, el grado de implementación de las medidas de control y prevención de la PPCCA, y un análisis de los resultados de la política para cada una de las

autoridades ambientales regionales seleccionadas (p. 63). Esto se hizo con el propósito de comparar el desempeño de la política entre las distintas zonas de estudio.

El segundo componente consistió en el análisis de resultados, en donde se debía evaluar los efectos de la PPCCA sobre los niveles de contaminantes criterio[65] en las zonas de las autoridades ambientales seleccionadas, y analizar la correlación entre las medidas implementadas en el marco de la política (p. 63). Luego, a partir de las correlaciones realizadas, se debió construir simulaciones de escenarios prospectivos, en donde se estimaba el comportamiento de las variables de resultado a partir de diferentes niveles de implementación de las medidas de control y prevención (p. 64).

Por último, el tercer componente brindó lineamientos para la construcción de las recomendaciones, los cuales las orientaron a mejorar la eficacia de las disposiciones de la política y a incorporar nuevas medidas de prevención y control para reducir los niveles de contaminantes criterio (p. 64). Los objetivos específicos fueron los siguientes (p. 3):

- “Analizar el nivel de implementación de las medidas establecidas en el marco de la Política de Prevención y Control de la Contaminación del Aire y su efecto sobre los indicadores de resultados, en la jurisdicción de las autoridades ambientales seleccionadas por el DNP para el periodo 2010-2015”.
- “Evaluar el efecto de los resultados de las medidas de prevención y control sobre el desempeño y la evolución de los niveles de los contaminantes criterio en la jurisdicción de las autoridades ambientales seleccionadas por el DNP, para el periodo 2010-2015”.
- “Analizar la correlación entre las medidas implementadas en el marco de la Política de Prevención y Control de la Contaminación del Aire, los resultados observados y los efectos sobre los niveles de contaminantes criterio, en cada jurisdicción de las autoridades ambientales seleccionadas para el periodo 2010-2015”.
- “Plantear diferentes escenarios de simulación prospectiva para estimar el comportamiento de las variables de resultados y de los niveles de contaminantes criterio, a partir de diversos niveles de implementación de medidas para prevenir y controlar la contaminación del aire”.
- “Generar recomendaciones sobre las disposiciones planteadas por la Política, así como proponer nuevas medidas, que permitan mejorar su efectividad en cuanto a sus resultados y su efecto sobre los niveles de contaminantes en la atmósfera”.

[65] Se refiere a: Material Particulado (MP), Óxidos de Azufre (SOx), Óxidos de Nitrógeno (NOx), Monóxido de Carbono (CO) y Ozono (O3).

Tal como se puede observar, los objetivos específicos tienen coherencia con el alcance descrito en los componentes de la evaluación. En términos de satisfacción de los estándares que corresponden a reglas de alcance, todo lo anterior abordó de manera adecuada los estándares de finalidad y fundamentación de la evaluación, objeto y alcance, y objetivos específicos.

Al igual que en los demás procesos de evaluación, los componentes del estándar de oportunidad, pertinencia y utilización que corresponden a reglas de alcance fueron incluidos en los lineamientos del pliego de condiciones relacionados con el PTIR. Para su construcción se requirió que debía incluir: metodología para identificar y priorizar las recomendaciones de la evaluación; la relación entre las recomendaciones y los objetivos del programa; la relación entre la evidencia de la evaluación y las recomendaciones; la justificación de cada recomendación; identificación de actores clave a los cuales afectarían las recomendaciones (p. 73). Todos estos lineamientos fueron abordados de manera satisfactoria en el documento PTIR, por lo que este estándar también fue satisfecho. La única evidencia que se encontró acerca del estándar de desarrollo de capacidad fue la pregunta orientadora "¿Cuáles indicadores permiten medir el avance en los resultados de la Política de Prevención y Control de la Contaminación del Aire?" (p. 63), lo cual es suficiente para que pueda considerarse como parcialmente satisfecho.

10.2.2 Reglas de frontera y participantes

Dentro del proceso de evaluación se incluyeron diversos actores con interés en la política. Particularmente, en la etapa de construcción de las recomendaciones se realizó una mesa de expertos en la cual participaron: MADS, DSEPP, Ecopetrol, Universidad de la Salle, ANDI, Caia Ingeniería, y Economía Urbana (DSEPP-DNP, 2017d, p. 3). Sin embargo, dentro del Comité de Diseño y en el Comité de Seguimiento a parte del DNP, únicamente participó el MADS[66]. Teniendo en consideración que en el CONPES 3943 de 2018 se establece como necesaria una articulación interinstitucional para el abordaje de la problemática de la política, y que varias de las acciones del plan de acción de la PPCCA formulada en 2010 establecía como responsables a otros ministerios. Incluir en estas etapas únicamente a MADS implica una satisfacción parcial de los estándares de participación de los actores relevantes, enfoque asociativo y consideración de evaluación conjunta.

Por parte del proceso de selección del equipo evaluador, se adelantó una modalidad de Concurso de Méritos Abierto (DSEPP-DNP, 2016f, p. 7), a través del portal SECOP. En el aviso de convocatoria se aclara que las propuestas fueron recibidas hasta el 28 de octubre de 2016 (DNP, 2016d, p. 1). Los proponentes pasaron por

[66] Esto con base en la respuesta a derecho de petición del DNP N° 20193600428271.

una etapa de habilitación a partir de criterios de capacidad jurídica, técnica, financiera y organizacional (p. 4), para luego proceder a la etapa de calificación de las propuestas a partir de: desarrollo metodológico, estrategia de recolección de información, experiencia y publicaciones del equipo de trabajo, experiencia adicional del proponente y apoyo a la industria nacional (p. 5).

El perfil del equipo de evaluación requerido consistió en un director del proyecto con título mínimo a nivel de maestría, con experiencia profesional de 3 año en dirección, diseño y/o implementación de estudios o consultorías relacionados con gestión o control ambiental; un experto en calidad ambiental con título mínimo a nivel de maestría con experiencia profesional de dos años en políticas y/o proyectos relacionados con calidad del aire, fuentes fijas y/o fuentes móviles; y un experto cualitativo con título de posgrado mínimo a nivel de maestría con experiencia profesional de dos años en diseño y/o implementación de técnicas cualitativas para el análisis de políticas o programas (p. 68).

Los proponentes fueron: Economía Urbana, Unión Temporal entre el Centro Nacional de Consultoría y R&R Ingeniería LTDA; y CPC Oriente. De estos, únicamente los dos primeros resultaron habilitados (Resolución 4386 de 2016, p. 4). Posteriormente, ganaría Economía Urbana Limitada a partir de obtener 31/40 puntos en la calificación del enfoque metodológico, en contraposición a 12/40 de la Unión Temporal. Todo lo anterior satisfizo el estándar de equipo de la evaluación. Finalmente, el valor de la consultoría pagado por parte del DNP fue de $383.971.600 (p. 4), y dentro de las responsabilidades contractuales del consultor se consideró poner a disposición el equipo de trabajo ofertado con los requerimientos técnicos de la consultoría (DSEPP-DNP, 2016f, p. 4), por lo que el estándar de recurso también fue satisfecho.

10.2.3 Reglas de elección y acciones

En cuanto a las acciones de los participantes y sus reglas, en primer lugar, el análisis de la evaluación fue desarrollado a partir de la elaboración del pliego de condiciones definitivo, y desde los requerimientos que en este documento se le efectuaron al consultor. En particular, se requirió dentro de la propuesta metodológica que se debía demostrar claridad conceptual y conocimiento sobre la PPCCA comenzando con la descripción detallada de la cadena de valor, el marco regulatorio, y la identificación, clasificación y descripción de las medidas de prevención y control de la política (p. 28).

También se analizó la disponibilidad de información en las zonas a seleccionar como casos (p. 63), que se contaba con una adecuada reconstrucción de la lógica de la intervención y de sus indicadores desde el documento de política del 2010, y que se analizaron los distintos riesgos que pudieron afectar el equilibro contractual como: Errores en el proceso de evaluación de las propuestas (p. 51), Incumplimiento de las obligaciones contractuales (p. 53); o la Alta rotación del personal que

compone el equipo mínimo de trabajo (p. 54). De esta manera se satisfizo el estándar de evaluabilidad.

Adicionalmente se puede observar que el proceso de evaluación tomó en cuenta los planes, actividades y políticas nacionales y locales en la materia. En el anexo 1 del producto final de la evaluación, se construye una matriz en la que se aclaran las medidas de prevención y control de la contaminación del aire evaluadas (33 en total), en donde se aclara el tipo de medida que es (regulación, información, educación, ordenamiento, control y vigilancia, incentivos, etc.), si el ámbito es nacional, regional o local, y el su soporte normativo (Economía Urbana, 2017a, p. 213). También en el anexo 2 se aclara el marco normativo de la PPCCA. Se considera la relación entre la PPCCA con el Plan Nacional de Desarrollo de 2010-2014 (DSEPP-DNP, 2016f, p. 3). Con todo lo anterior se cumplió el estándar de coordinación y alineamiento.

Con respecto a la selección del enfoque y del método, si bien se le pidió a los proponentes realizar una formulación de la metodología a aplicar, su discrecionalidad fue considerablemente limitada por las peticiones del DNP y del Comité de Seguimiento. Dentro de los cambios que son requeridos al primer producto de la consultoría, correspondiente al informe metodológico, se solicita que los indicadores de resultados deben contar con información antes y después de la implementación de las medidas (DSEPP-DNP, 2016g, p. 4). Esto corresponde a un “before-and-after desing” en términos de Morra y Rist (2009, p. 298), por lo que se solicita de manera explícita un tipo de metodología en particular.

Como en el resto de las evaluaciones, este primer producto consistió en un perfeccionamiento de la primera propuesta metodológica calificada en el proceso de contratación, y parte de una serie de requerimientos explicitados en el pliego de condiciones definitivo. Dentro de los lineamientos metodológicos que fueron incluidos para el primer producto se incluyó: establecer un marco conceptual de la evaluación; establecer las variables y los indicadores; relacionar cada pregunta de la evaluación con las variables propuestas; metodología de análisis de fuentes secundarias; definición de instrumentos de captura de información y técnicas de recolección de datos; identificación de los factores que afectan la efectividad de las medidas sobre los indicadores (variables de control); correlaciones con modelos estadísticos o econométricos (DSEPP-DNP, 2016f, p. 65). En resumen, todos los componentes del estándar de selección de enfoque y método fueron adecuadamente direccionados.

Se formularon cinco preguntas para el primer componente, nueve preguntas para el segundo, y cinco preguntas para el tercero. Estas preguntas conservan coherencia con el alcance descrito y con los objetivos específicos de la evaluación. En consonancia con el tipo de evaluación que fue acordado en el Comité de Diseño, el principal criterio de evaluación que se incluye en las preguntas orientadoras corresponde al de efectividad. También se incluyeron preguntas con criterios de evaluación de pertinencia y sostenibilidad (pp. 63-64):

Efectividad:

- “¿Cuál ha sido la evolución de estos indicadores en el período 2010-2015?”
- “¿Qué factores han afectado la efectividad de las medidas sobre los resultados?”
- “¿Qué condiciones y/o factores comunes (ambientales, políticos, económicos y sociales) comparten zonas objeto de estudio con niveles similares de resultados?”
- “¿Cuál fue la evolución de los niveles de contaminantes criterio en el periodo 2010-2015? ¿Cómo se compara su desempeño frente a los estándares regulatorios nacionales y aquellos definidos por la Organización Mundial de la Salud (OMS)?”
- “¿Cómo se relacionan las medidas de prevención y control con los resultados?”
- “¿Cómo se relacionan los resultados con los niveles de contaminantes criterio?”
- “¿Cómo se relacionan las medidas implementadas con los resultados presentados en los niveles de contaminantes criterio?”
- “¿Cuáles medidas de la Política de Prevención y Control de la Contaminación del Aire han presentado mejores resultados?”
- “¿Cuáles medidas de prevención y control han resultado más efectivas sobre los niveles de contaminantes criterio?”
- “¿Cuáles condiciones o factores han afectado el desempeño en resultados y en los niveles de contaminantes criterio?”
- “¿Cuál es el comportamiento de las variables de resultados y de los niveles de contaminantes criterio para diferentes grados de implementación de medidas de prevención y control?”

Pertinencia:

- “¿Las medidas implementadas han sido suficientes para reducir o mantener los niveles de contaminantes criterio dentro de los máximos permitidos por la regulación (tanto nacional como la definida por la OMS)?”
- “¿Cuáles medidas de prevención y control deben ser reforzadas o modificadas?”
- “¿Cómo pueden ser implementadas estas medidas para mejorar su eficacia sobre los resultados y sobre los niveles de contaminantes criterio?”
- “¿Qué condiciones o factores locales se deben tener en cuenta en la implementación de las medidas de prevención y control?”
- “¿Qué medidas de prevención y control adicionales se deben implementar para mejorar los niveles de contaminantes criterio?”

Sostenibilidad:

- “¿Cuáles escenarios de proyección del comportamiento de las variables de resultados se pueden estimar a partir de diversos niveles de implementación de medidas para prevenir y controlar la contaminación del aire y sus efectos sobre los indicadores de resultados?”

A pesar de que no existen menciones puntuales a las preguntas orientadoras dentro del informe final de la evaluación, y de que no se incluyó una matriz o algún tipo de guía que ayude a encontrar las respuestas, en este informe de evaluación se responde a todas las preguntas orientadoras. En este orden de ideas, los estándares de cuestiones a responder, selección y aplicación de criterios de evaluación, y respuesta a preguntas evaluativas, fueron satisfechos completamente.

Adicionalmente, se puede observar el requerimiento de una “matriz de consistencia de la evaluación” dentro del pliego de condiciones, como uno de los lineamientos del primer producto (p. 69), con lo cual se debió establecer la relación entre los objetivos de la evaluación, las preguntas orientadoras, las variables de análisis, los indicadores, y las preguntas de los instrumentos de recolección de datos. Esto facilitó la presentación de los hallazgos de la evaluación en el informe final y en el PTIR. Particularmente en este último se puede observar la derivación lógica entre las recomendaciones y los hallazgos de la evaluación, a partir de la forma en que son presentados. Para cada recomendación se construyó una ficha en la que se presentó la relación de la recomendación con los objetivos de la política, los hallazgos que sustentan la recomendación, una descripción de la recomendación aclarando los posibles mecanismos para su implementación, y los actores que son afectados por la implementación de la recomendación (Economía Urbana, 2017b). Con esto, se puede establecer que el estándar de claridad del análisis fue satisfecho.

El resumen ejecutivo realizado no utiliza subtítulos y presenta los hallazgos que se consideraron más relevantes, por lo que no es un resumen que satisfaga los componentes del estándar de claridad y representatividad del resumen de la evaluación. Nuevamente este estándar se satisface a partir de la construcción de la ficha técnica de la evaluación, la cual tiene una breve descripción de la intervención evaluada, el objetivo de la evaluación, la metodología de la evaluación, los principales resultados y conclusiones y las recomendaciones formuladas.

En relación con el informe final, en este se realiza una adecuada descripción del contexto de la intervención. En el primer capítulo, destinado a la construcción de los casos a comparar, se incluyen descripciones del contexto institucional y de políticas (Economía Urbana, 2017a, pp. 14-24), contexto ambiental (pp. 25-31), condiciones económicas (pp. 32-34), y el contexto social (pp. 38-48). De esta manera, el estándar de contexto de la intervención fue completamente satisfecho.

En cuanto a la lógica de la intervención, se observa que a pesar de que requiere su reconstrucción desde el primer producto de la consultoría, y que se contaba con el documento de política pública de 2010, en donde se aclaraba el plan de acción, ni en el informe final ni en el PTIR se presenta una descripción de la lógica de la intervención. Es importante aclarar que la evaluación no abordó toda la PPCCA, sino únicamente aquellas actividades de la política que estuvieron orientadas a la prevención o control de la contaminación del aire por contaminantes criterio (Economía Urbana, 2017a, p. 1). Estas comprenden tanto las que se incluyeron en el documento del 2010, como las que se desarrollaron posteriormente en la normatividad vigente (p. 11), y se realiza un listado de estas medidas, pero no se aclara su lógica. En este sentido, el estándar de lógica de la intervención no fue satisfecho en este proceso de evaluación.

En contraposición a la explicación de la metodología para la construcción de las recomendaciones que se incluyó en el PTIR, en el informe final no se incluye una sección particular del documento para la explicación de la metodología de la evaluación. Aun así, se manifiesta que la aproximación metodológica cuantitativa incluyó la construcción de indicadores de resultado, análisis estadístico del comportamiento de las variables, y la construcción de modelos para la estimación de reducción de emisiones (Economía Urbana, 2017a, pp. 12-13). Con respecto al análisis cualitativo, únicamente se manifiesta que este correspondió al análisis de contenido y estrategias con teoría fundamentada (p. 13). Por su grado de simplicidad, únicamente puede considerarse como parcialmente satisfecho el estándar de explicación de la metodología. En contraste, se incluyó un anexo particular titulado “Omisiones dentro de los análisis” (p. 249), en el cual se aclara la información que no se incluyó en los análisis estadísticos por presentar comportamientos atípicos de las variables, lo que satisfizo el estándar de mención de cambios y límites de la metodología de evaluación.

Con respecto a las fuentes de información, se manifestó la realización de 50 entrevistas semiestructuradas a personal de entidades del nivel regional y nacional –MADS, Min. Transporte, autoridades ambientales, secretarías, agremiaciones y expertos- (p. 1), para el análisis cualitativo. Como fuentes de información cuantitativa se seleccionó seis áreas geográficas, con 28 municipios (pp. 11-12) para la medición de resultados, y se recolectó información de los registros de las estaciones de monitoreo de calidad del aire, los reportes de emisiones de fuentes fijas del Registro Único Ambiental (RUA), inventarios de información de las autoridades ambientales, el Registro Único Nacional de Transito (RUNT), e información del IDEAM. Todo lo anterior direccionó de manera adecuada el estándar de validez y fiabilidad de las fuentes de información.

La recolección de datos cualitativos permitió que se considerara en la evaluación, la perspectiva de todos los actores relevantes. Asimismo, dentro del pliego de condiciones definitivo, se estableció como responsabilidades del consultor, la protección de los derechos y los datos personales de los actores consultados (p. 5),

mientras que al DNP le correspondió anonimizar la información producto de la consultoría que pudiera tener relación con el derecho a la intimidad, habeas data, y buen nombre (p. 6). Todo esto satisfizo el estándar de consulta y protección de implicados.

No se encontró evidencia de menciones de discrepancias entre miembros del equipo evaluador, por lo que este estándar no se satisfizo. Finalmente, dentro del contrato de consultoría (DNP, 2016e) se designó como fecha de entrega del informe final el 30 de mayo 2017, y del PTIR el 4 de julio del mismo año. Estos documentos presentan exactamente estas fechas de elaboración. Adicionalmente en los conceptos a los productos realizados por la DSEPP se confirma que fueron entregados en estas fechas (DSEPP-DNP, 2017d, p. 5; 2017e, p. 5). Lo cual significa que se cumplió con el estándar de evaluación dentro del plazo y el presupuesto asignados.

10.2.4 Reglas de agregación y control

Al igual que en el resto de los contratos, hubo un control de calidad continuo a lo largo de todo el proceso de evaluación. Este fue canalizado a través de distintos mecanismos: las responsabilidades del consultor incluyeron entregar los informes de avance que le solicitara el supervisor del contrato -DSEPP- (DSEPP-DNP, 2016f, p. 5); el DNP tuvo como competencia ejercer el control del cumplimiento del contrato y expedir el certificado a satisfacción de los productos (p. 6); garantía de cumplimiento de las obligaciones del consultor a favor del DNP a partir de contrato de seguro en póliza, patrimonio autónomo o garantía bancaria (Contrato, p. 9); cláusula penal compensatoria (p. 10); y cláusula de multas. Las cuales cubren riesgos de incumplimiento, baja calidad del servicio y pago de salarios y prestaciones sociales.

También se estableció como responsabilidad del Comité Técnico de Seguimiento la elaboración de comentarios a los productos, y de la Secretaría Técnica del Comité, el recibir, consolidar, depurar y dar el tratamiento respecto a dichos comentarios (DSEPP-DNP, 2016f, p. 60), en efecto, los conceptos técnicos realizados a los productos de la consultoría por parte de la DSEPP, describen el proceso de inclusión de dichos comentarios en los productos para la elaboración de recibo a satisfacción. Todo lo anterior garantizó el cumplimiento de los estándares de control de calidad y de inclusión de observaciones de los implicados.

10.2.5 Reglas de posición

Frente a los estándares de independencia del evaluador y del proceso evaluativo libre y abierto, en el contrato de consultoría fue incluida la cláusula correspondiente a la autonomía del evaluador (p. 9), en la cual se establece que este actuaría con total autonomía técnica y administrativa en el cumplimiento de sus obligaciones.

También es reconocido esto por parte de los funcionarios en el MADS, los cuales al ser cuestionados por la respuesta que tuvo el equipo consultor a sus demandas puntualizaron: "… no es que respondieran a nuestras demandas, porque eran totalmente autónomos, pero sí nos tuvieron al tanto de cómo iba el proceso" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

Adicionalmente, en el concepto del producto 5 se aclaró lo siguiente ante la eliminación de una de las recomendaciones propuestas por el consultor, pero que fue eliminada en la mesa de expertos: "…si la firma lo considera, lo puede incorporar al PTIR con el sustento adecuado. La mesa de expertos no es camisa de fuerza, sino instrumento de validación. El criterio final corresponde a la firma" (DSEPP-DNP, 2017d, p. 3). Sin embargo, existe evidencia que va en dirección opuesta.

En los conceptos técnicos de los productos de la consultoría se describe el proceso de transformación de estos a partir de los comentarios realizados por el Comité de Seguimiento. En esta descripción se puede observar, que a pesar de la cláusula de autonomía y de su reconocimiento, la estructura institucional de control de calidad restringe la independencia del evaluador. En el concepto del primer producto se restringe el uso de formularios con el objetivo de recolectar información cuantitativa (DSEPP-DNP, 2016g, p. 4); en el concepto del producto 4, se observan comentarios direccionados a la modificación de las valoraciones realizadas por el evaluador, justificadas en la no consideración de determinadas variables como la pluviosidad, información mediática, recursos humanos que entran a través de convenios, condiciones meteorológicas, y diferencias de medidas entre distintas instituciones (DSEPP_DNP, 2017e, pp. 4-5).

Información más significativa se encuentra en el concepto técnico del PTIR (DSEPP-DNP, 2017d), en el que se requieren cambios de naturaleza substancial, que pueden considerarse en contravía de la autonomía técnica del evaluador al requerirse para proceder con el pago al consultor. Entre estos requerimientos están: evaluar la pertinencia de medir otros contaminantes no criterio, pero que son importantes por área geográfica (p. 3), resaltar que el éxito de la política depende del involucramiento de los sectores generadores de contaminantes -como transporte y minas-, presentar un análisis de la ineficacia y poca efectividad de la coordinación a través de los COTSA -Consejos Territoriales de Salud Ambiental- (p. 5), la no inclusión de la modificación del Decreto 2872, la inclusión como recomendación de la realización de un CONPES, y la redefinición de los contaminantes criterio. En este sentido, no se cumplen los estándares de independencia de los evaluadores, ni de proceso evaluativo libre y abierto.

10.2.6 Reglas de información

Frente a los estándares correspondientes a reglas de información, se tiene que el pliego de condiciones definitivo, como producto del Comité de Diseño de la

evaluación, cumplió con todos los componentes del estándar de documentación acerca de la finalidad y la expectativa. Existieron lineamientos dentro de esta documentación relativos a la claridad del informe final, tales como: "… presentar brevemente la metodología utilizada para dar respuesta a cada una de las preguntas orientadoras y los resultados de la evaluación" (DSEPP-DNP, 2016f, p. 72); "… presentar las conclusiones de la evaluación, con base en los resultados para las variables de análisis utilizadas y las preguntas orientadoras" (p. 72). Ninguno de estos dos lineamientos fue cumplido en el informe final de la evaluación, razón por la que el estándar de informe de la evaluación no se considera satisfecho.

Finalmente, con respecto a la difusión de la evaluación, se consideró la presentación de los resultados ante los directivos de las entidades que conformaron el Comité de seguimiento (p. 5). En efecto, en la entrevista realizada al MADS, se comentó al respecto: "… ese documento se está socializando en diferentes mesas con diferentes actores para poder tener la retroalimentación y llegar finalmente a un documento definitivo de política pública" (Coordinador Grupo de Gestión Ambiental Urbana, MADS). Se hizo también la publicación del resumen ejecutivo, la ficha técnica y el informe final en el portal de SINERGIA. El estándar de difusión fue satisfecho de esa manera.

10.3 Análisis de los patrones de interacción de los participantes en el proceso de evaluación de la PPCCA

Con relación a los patrones de interacción de las situaciones de acción del proceso de evaluación, se encontró que a pesar de la ausencia de actores importantes a los cuales se les asigna responsabilidades en la implementación de las recomendaciones, todas estas tienen como responsable –por lo menos parcial- al MADS. De la misma forma, en el PTIR se puede verificar que de las cuatro recomendaciones que fueron implementadas, dos incluyen como responsables al Ministerio de Minas y al Ministerio de Transporte, sin embargo, estas entidades estaban desarrollando un proceso paralelo de coordinación en la CONASA con el MADS, para la formulación del nuevo CONPES. Esto significa que, a pesar de la ausencia de estas entidades en los comités de la evaluación, el MADS tuvo la suficiente influencia para introducir las recomendaciones en la política, por lo que el involucramiento de usuarios previstos fue de un nivel medio alto.

La relación que se estableció entre el MADS, el DNP y el equipo evaluador de Economía Urbana se caracterizó por un adecuado grado de reactividad frente a las preocupaciones y demandas de los usuarios previstos. Al ser preguntados acerca de la coordinación que se estableció entre el Comité de Seguimiento y el equipo consultor, en el MADS se comentó:

> "… parte del éxito de esto fue que participamos en la formulación de esta consultoría de evaluación de la política. La participación activa del MADS dentro de la formulación y dentro de la implementación de esta evaluación llevó a que los resultados fueran 100% útiles para lo que nosotros

> necesitamos (…) fue un acercamiento total en el cual ellos nos dieron la oportunidad de expresarles todos los avances que hemos tenido, de darles las evidencias de cómo íbamos, por qué sí o por qué no, y de determinar cuáles eran los puntos críticos en que considerábamos debería enfocarse más en la evaluación de la política. Y aunque algunos de ellos los tomaron y los priorizaron mientras que otros no los consideraron nos dieron esa oportunidad; porque a veces se llega la evaluación de la política y no se tiene en cuenta esto y se hace una evaluación totalmente externa, con información externa, y nos reconoce ciertos procesos que se han adelantado, que, si bien son exitosos o no, dan un matiz diferente a la evaluación (…) Entonces, la vinculación del ministerio en la evaluación desde el comienzo creo que fue el factor determinante del éxito de este producto que se generó" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

Tal como se observa, se tiene una percepción positiva acerca del grado de incidencia que tuvo el MADS en la definición del diseño de la política, y acerca del grado en que el proceso de evaluación respondió a sus necesidades de información, y de su participación en la definición de los temas a los cuales iría enfocada la evaluación. Esto refleja reactividad particularmente en el tema de definición de las opciones de preguntas de la evaluación. También se observa un proceso de trabajo conjunto para la obtención de un acuerdo alrededor del alcance de la evaluación:

> "Al comienzo llegamos con la expectativa de que podríamos evaluar el impacto que ha tenido la política en la calidad del aire. Pero al entender los niveles de profundidad que se le pueden dar a una evaluación comprendimos que no teníamos la suficiente información para poder evaluar eso. Entonces llegamos con DNP a decir que se iba a hacer una evaluación de implementación para ver el nivel en que se ha implementado y qué acciones se han tomado, y con eso quedamos totalmente satisfechos" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

Al no medir el impacto, se restringieron las mediciones de la evaluación a la valoración de los niveles de contaminantes criterio en el aire, y se descartó hacer mediciones de variables acerca de la salud en la población. A pesar de que el alcance de la evaluación no correspondió a una "evaluación de implementación", sino de resultados, se observa que los usuarios previstos quedaron satisfechos con los acuerdos logrados a partir de las deliberaciones y del análisis de viabilidad de la evaluación que fue realizado en la etapa de diseño de la evaluación.

Respecto a la coordinación que se dio a través de debates metodológicos, la mejor información al respecto se encuentra en los comentarios que el Comité de Seguimiento realizó a los productos. Estos comentarios se encuentran en los conceptos técnicos que realizó la DSEPP a dichos productos. Para el producto correspondiente al informe metodológico, el Comité de Seguimiento realizó comentarios el 22 de diciembre de 2016, y el 26 de diciembre se recibió el producto

ajustado por parte del equipo evaluador (DSEPP-DNP, 2016g, p. 2), razón por la cual se puede considerar que no hubo mayor resistencia por parte del evaluador en la incorporación de los comentarios relacionados con la metodología. Los comentarios incluidos en el informe metodológico por parte del equipo evaluador y que destacan la reactividad del evaluador a las preocupaciones y demandas de los usuarios, son los siguientes (p. 2):

- Usar “resultados” para hacer referencia en los resultados inmediatos, específicamente a cambios en las fuentes de emisión.
- Replantear indicadores de producto y resultados aclarando: periodicidad, fuente de información, fórmula de cálculo y relevancia.
- Garantizar que los indicadores cuentan con información antes y después de la implementación de las medidas.
- Evitar el uso de encuesta para recolectar datos cuantitativos -usar datos disponibles en las entidades-.
- No evaluar resultados en salud.
- No dar tanta relevancia al levantamiento de información cuantitativa.

Al cuarto producto, correspondiente al informe final de la evaluación, el Comité de Seguimiento realizó los siguientes comentarios en temas metodológicos (DSEPP-DNP, 2017e, p. 6):

- Cada argumento debe tener evidencia que lo respalde.
- Establecer y describir los elementos y relaciones de un modelo integrado que sustente el análisis y la evaluación, con el que se pueda generar escenarios prospectivos concluyentes.
- Fórmula de cálculo y fuente de información para cada uno de los nuevos indicadores que se agregaron al producto.

También se menciona en el mismo documento, que todos estos comentarios del Comité de Seguimiento fueron incluidos por parte del equipo evaluador. La naturaleza de los comentarios, así como el hecho de que fueran incluidos en los productos refleja que también hubo reactividad en términos de debates metodológicos y adaptación de la metodología a lo largo del proceso de evaluación. Con relación al involucramiento de los usuarios previstos en la interpretación de los hallazgos y en la construcción de las recomendaciones, la evidencia más relevante que permite medir este punto se encuentra en el PTIR, en donde, tal como se comentó en los estándares de reglas de frontera, se incluyó a los funcionarios del MADS dentro de la mesa de expertos que sirvió como método de validación y priorización de las recomendaciones.

Las recomendaciones surgieron de un proceso que pasó por el análisis y discusión conjunta del equipo evaluador (Economía Urbana, 2017b, p. 6), luego se procedió a la mencionada mesa de expertos. Para esta mesa, el equipo evaluador elaboró una ficha con el fin de que los expertos evaluaran de manera rápida tres aspectos de cada recomendación -viabilidad de expedición, viabilidad de implementación e impacto en el mejoramiento de la calidad del aire- (p. 8). Luego se procedió a abrir

la discusión en donde se buscó recoger las percepciones positivas y negativas acerca de las recomendaciones, y sugerencias para su mejoramiento (p. 9), de donde surgió, adicionalmente, una nueva propuesta para incluir en el PTIR. Se aclara en el mismo documento que las sugerencias que surgieron de la mesa de expertos fueron tenidas en cuenta para la elaboración de este (p. 9). Como se observa, se desenvolvió una apropiada coordinación para la construcción de las recomendaciones de la evaluación. No se encontró ninguna evidencia acerca de planeación conjunta de seguimiento a la implementación de los resultados de la evaluación, o de búsqueda proactiva del uso por parte del equipo evaluador.

10.4 Análisis de la efectividad de la evaluación del proceso de evaluación de las medidas de defensa comercial de salvaguardia y antidumping

El proceso de evaluación produjo 12 recomendaciones. La primera consistió en fortalecer la Comisión Técnica Nacional intersectorial para la Salud Ambiental (CONASA) y los Consejos Territoriales de Salud Ambiental (COTSA), para esto, se sugiere que el MADS solicité a la secretaría técnica del CONPES, la elaboración de un documento CONPES para el ajuste de la PPCCA; que este CONPES sea formulado a partir de un proceso de coordinación intersectorial a través de la CONASA; e incluir en el Plan Nacional de Desarrollo 2018-2022 un artículo que impulse la gestión integral de la calidad del aire (Economía Urbana, 2017b, pp. 17-18). En relación con su implementación, en el MADS se comentó:

> "… esta parte de la coordinación institucional fue el punto clave que nos dijo que lo que necesitamos es un documento de política de otro nivel jerárquico. Y ahí pensamos en el documento CONPES (…) llevamos meses de trabajo con ellos [con los otros ministerios y entidades que firmaron el CONPES 3942 de 2018]. Este hallazgo que nos hizo en cierta forma el DNP, nos llevó a trabajar conjuntamente en ese documento CONPES, y ahí ya todos firman el documento, todos son proponentes. Es importante decir que esta propuesta de elaborar el documento CONPES la hicieron tres ministros, eso nunca se había dado antes. Tres ministros firmaron un oficio hacia el DNP diciendo que el tema de calidad del aire no solo es de ambiente, sino que es de minas y de transporte también, y por eso era necesario hacer un documento CONPES entre todos" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

En efecto, el CONPES 3943 de 2018 fue firmado por DNP, Ministerio de Salud y Protección Social, Ministerio de Minas y Energía, Ministerio de Vivienda, Ciudad y Territorio y Ministerio de Transporte. Adicionalmente, en el documento *Bases para*

el Plan Nacional de Desarrollo 2018-2022[67]*,* Se incluye dentro de sus objetivos "mejorar la calidad del aire, del agua y del suelo para la prevención de los impactos en la salud pública y la reducción de las desigualdades relacionadas con el acceso a recursos" (p. 397), para el cual se establecen 3 líneas de acción en concordancia con el CONPES.

A pesar de lo anterior, en la cita se puede observar que la recomendación no fue una propuesta que surgiera del equipo evaluador, sino que fue una acción que ya se había entendido como necesaria antes de los hallazgos de la evaluación, por parte de las entidades de la CONASA. Esto se confirma a partir de los comentarios realizados por parte del Comité de Seguimiento al producto PTIR, en donde se manifestó expresamente: "Esto es una acción más enfocada al CONPES que a la CONASA. La idea es actualizar la política de prevención y control de la contaminación del aire a través de un documento CONPES" (DSEPP-DNP, 2017d, p. 7). En este sentido, no se da ningún aprendizaje o cambio de modelo mental de las entidades a cargo de la PPCCA, la recomendación es una justificación de una decisión previa, por lo que la implementación de esta recomendación debe descartarse como ejemplo de efectividad del proceso de evaluación.

La segunda recomendación consistió en incluir en la PPCCA lineamiento para diferenciar la implementación de las medidas de prevención y control según características geográficas y cumplimiento de la norma nacional (Economía Urbana, 2017b, p. 20). Para esto, se propuso establecer orientaciones sobre el tipo de monitoreo y tipos de medidas que se deberían implementar por parte de las autoridades ambientales de acuerdo con características específicas de las jurisdicciones (p. 21). Al respecto se respondió:

> "Eso es más de la gestión de las autoridades ambientales. El CONPES es más global y no llega a ese nivel de acciones dentro de las entidades territoriales. Entonces, sí estamos tratando de promover el tema de prevención y control pero sobre todo en los planes de descontaminación de las ciudades; en Medellín, en Bogotá, y en Calí, estamos tratando, a través de mesa de trabajo con ellos, de separar esas dos cosas de prevención y control" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

Lo anterior significa que a pesar de que se han implementado acciones en el mismo sentido que la recomendación, estas han estado fragmentadas según los lugares en los cuales se han implementado. La recomendación en el sentido que es propuesto por el equipo evaluador no es implementada. En tercer lugar, se recomendó priorizar medidas de política en contaminantes críticos y establecer normas de calidad del aire a partir de análisis de costo beneficio. Para esto, se sugirió flexibilizar el seguimiento a las emisiones de fuentes contaminantes en

[67] Este documento se encuentra en: https://colaboracion.dnp.gov.co/CDT/Prensa/ PND-2018-2022.pdf

función de la calidad del aire; redefinir los contaminantes y enfocar las acciones en aquellas medidas que permiten recudir las emisiones de material particulado y ozono, y no en medir óxidos de nitrógeno, monóxido de carbono y óxidos de azufre, puesto que estos generan costos innecesarios (Economía Urbana, 2017b, p. 24). En relación con su implementación se manifestó:

> "La recomendación de ellos de una nueva norma de calidad de aire basada en costo beneficio pues fue el trabajo que hicimos el año pasado, que es la Resolución 2254 del 2017. En esta norma se priorizan contaminantes, vuelve y se revisa cuáles son esos contaminantes con mayor efecto en salud, se incluyen unos que no estaban, y se toman metas al 2030. Aunque hay que resaltar que el tema de costo beneficio es uno de los grandes retos que tienen no solo los temas ambientales, sino muchos temas sociales del país. Paralelo a la evaluación de la política se hizo un estudio de costos de la degradación ambiental con apoyo del DNP que nos ayudó en a cuantificar el impacto. Pero los retos grandes que tenemos como gobierno, y en este proceso de ingresar a la OCDE, es poder cuantificar el tema del impacto y del costo de implementación y relacionar estos dos para llevar un trabajo totalmente armonizado en cuanto al tema del costo beneficio" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

En completa correspondencia, los contaminantes criterio y sus niveles permitidos fueron redefinidos y priorizados en el artículo 20 de la Resolución 2254 de 2017. El análisis costo beneficio fue parcialmente implementado, puesto que lo más importante es medir los beneficios en salud. En todo caso, esta recomendación surge a partir de los comentarios que el Comité de Seguimiento hizo al PTIR: "Por favor, ajustar e incluir, que los contaminantes criterio para Colombia deben ser nuevamente definidos, porque contaminantes como CO o PTS no son relevantes para ser medidos y si generan un costo de operación innecesario" (DSEPP-DNP, 2017d, p. 7). De nuevo, esta recomendación no significó un cambio en el modelo mental de los funcionarios del MADS, implica una justificación de una decisión previa, y debe ser descartada como ejemplo de efectividad del proceso de evaluación.

La cuarta recomendación fue desarrollar una socialización masiva regional y nacional del estado de la calidad del aire, los costos de salud, metas programadas de reducción y avances de los gobiernos, lo cual se sugirió a través de sistemas de información que mostraran en tiempo real las condiciones de calidad del aire, y poner a disposición de la ciudadanía toda la información referente a planes y proyectos orientados al mejoramiento del aire (Economía Urbana, 2017b, p. 28). Al respecto se contestó:

> "Nosotros estamos actualizando el sistema de reporte de información a la sociedad. Sin embargo, dentro de las recomendaciones que hacen ellos, es hablar de los avances de los gobiernos. Esto es un punto muy importante en

> el tema de evaluación de las políticas que es complejo, es ¿cómo, en el marco de la participación ciudadana, el ciudadano sabe qué están haciendo las diferentes regiones para mejorar la calidad del aire? Es un tema de transparencia en las acciones, y, a veces, va un poco más allá del Ministerio de Ambiente. Por ahora hemos avanzado en el tema de informar a la comunidad en tiempo real del estado de la calidad del aire y nos queda pendiente un poco ese tema de informar las acciones que ha tomado el gobierno, el tema de indicadores para que la comunidad sepa "ah, el ministerio, la secretaria distrital de ambiente, etc. se comprometieron en esto, entonces ¿cómo han avanzado en esos compromisos?". Y es un proyecto que quisiéramos avanzar bastante y se llama el tema de gobernanza de la calidad del aire" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

Estos avances en socialización y divulgación de los resultados fueron incluidos en la Resolución 2254 de 2017 (artículos 22, 23, 24 y 25), en donde se establecen requerimientos a las autoridades ambientales del diseño e implementación de una estrategia de comunicación para los niveles departamental, municipal y distrital; la obligación al IDEAM de realizar informes anuales del estado de la calidad del aire; e informes mensuales, trimestrales y anuales de la calidad del aire a través de páginas Web a las autoridades ambientales. En este sentido, la recomendación fue implementada y refleja efectividad.

En quinto lugar, se recomendó ampliar y modificar la regulación de fuentes fijas basada en permisos de emisiones. Particularmente se sugirió complementar el permiso de emisión con el instrumento de registro de emisión, de tal manera que fuese obligatorio para todas las fuentes fijas de emisión (Economía Urbana, 2017b, p. 31). Al respecto, únicamente se comentó: "El ministerio ya hizo un estudio que nos dio algunos soportes técnicos y estamos en el proceso de formulación de esas normas" (Coordinador Grupo de Gestión Ambiental Urbana, MADS). A julio de 2019, en la página del IDEAM se aclara que la normatividad sobre emisiones de fuentes fijas sigue siendo la Resolución 909 de 2008 y el *Protocolo para el control y vigilancia de la contaminación atmosférica generada por fuentes fijas* construido en el 2010. A pesar de que se estén dando acciones encaminadas en ese sentido, la recomendación no ha sido implementada y no refleja efectividad.

La sexta recomendación fue reglamentar instrumentos económicos para alcanzar las metas de reducción de emisiones donde se incumplen las normas y las fuentes fijas que contribuyen de modo significativo a la problemática del aire. Para esto, se sugirió reglamentar la tasa por emisión para que pueda ser cobrada a las fuentes digas que incumplan la norma nacional (Economía Urbana, 2017b, p. 35). Ante esto se respondió:

> "Nos encontramos también avanzando en el tema de instrumentos económicos. Lo importante ahí es que ya existe un impuesto al carbón. Eso es el tema de emisiones contaminantes que están relacionadas también con las emisiones de gases de efecto invernadero. Y parte también de recomendaciones que nos han hecho actores externos es no volver a generar

> doble tasación o doble requerimientos económicos a las industrias. Si ya se cobra un impuesto al carbón no hay que medir cuál es el beneficio que eso tiene en reducción de contaminantes como material particulada. Pero estamos avanzando en una tasa de emisiones por contaminantes al aire, que es parte de lo que también está en el CONPES" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

En el CONPES 3943 se pudo constatar la inclusión como línea de acción, la adopción de instrumentos económicos como tasas retributivas por emisiones contaminantes (p. 71). De acuerdo con esto, la recomendación fue implementada en los términos de la recomendación de la evaluación. La séptima recomendación estuvo relacionada con la anterior. Consistió en alinear los incentivos tributarios para fuentes fijas con programas de conversión tecnológica, cuyos mecanismos de implementación fueron: simplificar el trámite de los incentivos, y alinear el acceso de los incentivos a esfuerzos de reconversión tecnológica (Economía Urbana, 2017b, p. 38). Sobre la implementación de esta recomendación en el MADS comentan:

> "Ahí el punto es que ya tenemos unos incentivos tributarios para fuentes fijas. El inconveniente que tenemos es que no todas las personas están accediendo a esos incentivos, no todos están solicitando esos incentivos. Entonces el DNP lo que nos recomienda es facilitar el proceso de los incentivos tributarios. No obstante, ya que esto también es un proceso con el Ministerio de Comercio, parte del trabajo de la primera recomendación que es articulación entre entidades" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

En este sentido, para que la recomendación sea implementada se necesita primero avanzar en la coordinación interinstitucional de la PPCCA. En este caso, se puede argumentar que una de las razones por las que la recomendación no ha sido implementada, es la ausencia de MinCIT en el Comité de Seguimiento, el cual pudo ser el espacio de coordinación interinstitucional en el que esta entidad se hubiese podido apropiar de esta propuesta.

La octava recomendación fue racionalizar las obligaciones en monitoreo de calidad del aire y emisiones a los responsables de fuentes fijas, para lo que se requirió no obligar a los responsables de las fuentes fijas a realizar mediciones de la calidad del aire con excepción de las fuentes que por su magnitud pueden cambiar los niveles de calidad del aire. En los otros casos, la responsabilidad debe ser de la autoridad ambiental (Economía Urbana, 2017b, p. 41). Frente a la indagación por su implementación se comentó:

> "… una medida es no pedir tanto el monitoreo sino invertir más en las acciones de reducción. Pero en nuestro proceso de seguimiento, control y fiscalización a estas entidades es muy importante el monitoreo. Entonces hay

> que evaluar muy bien esa recomendación (..) Ellos lo que nos dicen es: 'el país está invirtiendo mucho en monitorear, en medir, y no en tomar inversiones en reducir, entonces reduzca el monitoreo y esa plata la puede invertir en acciones de reducción'. Sin embargo, a lo que nos lleva el monitoreo es a exigirle a las empresas que hagan acciones, por lo tanto a veces no es tan sencillo. No dejar de monitorear sino exigir más las acciones es un poco la visión que tenemos que evaluar como gobierno" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

En este sentido, la recomendación fue descartada por los usuarios previstos y, por lo tanto, no fue implementada. La novena recomendación fue realizar un etiquetado de vehículos para alinear las restricciones de movilidad con el nivel de emisiones de los vehículos. Este etiquetado consistiría en una clasificación de los vehículos según su nivel de contaminación determinado a partir de su tecnología y edad (Economía Urbana, 2017b, p. 45). Adicionalmente, se consideró necesario que este etiquetado permitiera el fácil reconocimiento y verificación, y que estuviera registrado en el RUNT. Al respecto, y teniendo en cuenta que para esto también se necesita avanzar en la coordinación con otro ministerio que no fue incluido en el Comité de Seguimiento, se comentó: "Sí, estamos trabajando en proyectos regulatorios en ese sentido (...) con el Ministerio de Transporte" (Coordinador Grupo de Gestión Ambiental Urbana, MADS). Sin embargo, a 2019 no se ha implementado esta recomendación.

La décima recomendación fue el establecimiento de impuesto de vehículos a partir de su edad y tecnología, de tal manera que se incentive la renovación de la flota. Para esto, se sugirió adicionar a la base gravable del valor comercial del vehículo, su desempeño ambiental a partir de las variables de edad y tipo de combustible utilizado, con lo que se establecerían mayores impuestos a vehículos de mayor edad y con tecnologías más contaminantes (Economía Urbana, 2017b, p. 49). Al respecto, se respondió lo siguiente:

> "… el impuesto a vehículo es un tema más de transporte y de comercio, y también de hacienda, entonces se sale del ministerio. Nosotros apoyamos, promovemos, pero es un trabajo que requiere esa articulación entre entidades (…) se deben crear mecanismos que desincentiven el uso de vehículos altamente contaminantes. Para ello, en 2019 el Ministerio de Transporte, con el apoyo del Ministerio de Ambiente y Desarrollo Sostenible, realizará un análisis técnico y formulará una propuesta de modificación a la ley que establece el impuesto de vehículos (Ley 488 de 1998). Lo anterior, en función de su rodamiento y externalidades generadas por la edad, como el riesgo de accidentalidad, entre otros, de tal manera que se beneficien las tecnologías vehiculares de cero y bajas emisiones. En el mismo año, el Ministerio de Transporte presentará la propuesta al Congreso de la República" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

La normatividad alrededor del impuesto a vehículos está determinada en el artículo 145 del Ley 488 de 1998, y sus tarifas son actualizadas cada año a través de un

decreto del Ministerio de Hacienda. Para el año 2019 no se evidencia ninguna modificación de este artículo, y el Decreto 2439 de 2018 del Ministerio de Hacienda, no estableció ningún lineamiento en la dirección de la recomendación. Esto significa que la recomendación no ha sido implementada.

En onceavo lugar, se recomendó la inclusión del análisis costo-beneficio ambiental en las decisiones de renovación de tecnologías vehiculares de transporte público masivo y transporte público integrado en las áreas urbanas. Esto incluiría la valoración económica de impactos ambientales de distintas alternativas tecnológicas y de combustible, informar a los tomadores de decisión sobre dichos efectos, e incorporar este análisis como requisito de la aprobación de la cofinanciación nacional de estos sistemas (Economía Urbana, 2017b, p. 52). Frente a la implementación de esta recomendación se respondió:

> "Esas decisiones son locales, son de las entidades de los municipios y de los grandes centros urbanos y en este caso pues tenemos todo el proceso de licitación de Transmilenio, donde hemos participado activamente tratando de incidir en que esta recomendación se adopte por parte de los tomadores de decisión, en este caso la Alcaldía. Entonces, ya el ministerio en diferentes ocasiones se ha comunicado vía oficios con el alcalde de Bogotá diciéndole que es importante que evalúe el tema de tecnologías vehiculares y promueva el transporte con cero o bajas emisiones en función del análisis costo beneficio. Es compleja la respuesta, porque ellos hablan de que no existen tecnologías y que los costos incrementarían demasiado la tarifa del transporte, pero pues se sabe solo cuando se haga ese análisis costo-beneficio" (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

A pesar de que esto no fue implementado en la renovación de la flota de Transmilenio en 2019, en el CONPES 3943 de 2018, en la línea de acción de renovación y modernización del parque automotor, se establece que, en el 2020, el Ministerio de Minas y Energía elaborará un estudio de costo-beneficio del impacto en la reducción de emisiones por el posible aumento en el uso de biocombustibles. Adicionalmente, se establece que el diseño de los programas que hagan parte de esta línea de acción deberá considerar el "análisis de costo beneficio de la implementación de las tecnologías vehiculares más limpias" (p. 58). Lo anterior significa que, a pesar de que en la entrevista no se resaltan acciones implementadas en esta dirección, el CONPES permite establecer que la recomendación si fue implementada.

La última recomendación fue establecer calendarios y metas de renovación del parque automotor diésel (transporte de carga y colectivo) en grandes ciudades, como pre-requisito a un suministro de diésel de mejor calidad. Para esto, se sugirió adicionalmente enfocar los incentivos tributarios en la adquisición de vehículos más limpios, y programas de desintegración de este tipo de vehículos en zonas urbanas (Economía Urbana, 2017b, p. 52). Al respecto solo se comentó: "También está

dentro del Conpes porque es un tema del ministerio de minas y energía de modernizar sus refinerías para mejor combustible” ” (Coordinador Grupo de Gestión Ambiental Urbana, MADS).

En efecto, en el CONPES se establece que se implementará una estrategia de renovación y modernización del parque automotor priorizando las categorías que son altamente contaminantes; se evaluarán los programas de renovación y reposición de camiones y buses; se definirá en 2020 la estrategia para aumentar la desintegración de camiones y buses de servicio público incluyendo metas y criterios técnicos; y se adoptará el acto administrativo correspondiente en 2021; por último, se espera que a 2024 el Ministerio de Transporte publique un informe con el avance de la estrategia. Con todo lo anterior, la recomendación si fue implementada en el sentido sugerido por la evaluación. En general, con respecto a su grado de uso se comentó:

> “… básicamente de ese proceso se definió la necesidad de generar una política de estado. Entonces, pasamos de esta Política de Prevención y Control de la Contaminación del Aire a ver la necesidad de obtener un documento CONPES. Eso fue un poco el resultado de la visión que nosotros le dimos. Actualmente estamos trabajando en ese documento CONPES, en el cual estamos tratando de integrar varias de las recomendaciones que nos dio la evaluación. Y lo estamos trabajando con el DNP, ellos dijeron "muy chevere este tema de planeación y tenemos que empezar articular otros actores que no sean solo del tema de ambiente (…) sí aportó, puesto que la evaluación en conjunto con las recomendaciones que nos dieron es el eje de lo que estamos trabajando, es el eje de la cadena de valor y de la programación del cuatrienio que viene, y es la base que le vamos a dar al nuevo gobierno que entre para trabajar. Entonces, creo que las recomendaciones están totalmente articuladas con la experiencia que hemos vivido desde el ministerio" (Coordinador Grupo de Gestión Ambiental Urbana, MADS)

Esto significa que la evaluación a esta política fue desarrollada en un contexto en el cual se deseaba medir los resultados de la política elaborada en 2010, y en donde se esperaba que sus resultados sirvieran como insumo para el nuevo proceso de formulación de política pública para el cuatrienio de 2018 a 2022, correspondiente al nuevo gobierno.

10.5 Conclusiones

En cuanto al cumplimiento de los estándares de calidad, en este proceso de evaluación se presentaron problemas en los estándares de participación de los actores relevantes, enfoque asociativo, consideración de evaluación conjunta, lógica de la intervención, explicación de la metodología en el informe final, mención de discrepancias entre miembros del equipo evaluador, independencia de los

evaluadores, proceso evaluativo libre y abierto y, por último, informe de la evaluación.

A pesar de que no fueron involucradas como usuarios previstos varias de las entidades de las que se asignan como responsables de la implementación de las recomendaciones, y que además hacen parte de la CONASA, el involucramiento del MADS fue decisivo. Se desenvolvieron patrones de interacción orientados a la coordinación efectiva entre el MADS, el DNP y el equipo evaluador. Se observa reactividad del equipo evaluador y del DNP frente a las preocupaciones y demandas de los usuarios en los temas de definición de preguntas, alcance de la evaluación, metodología, adaptación de la metodología, interpretación de hallazgos y construcción de recomendaciones.

De las recomendaciones que fueron formuladas por el equipo evaluador, que luego atravesaron por un proceso de perfeccionamiento a través de la mesa de expertos (en la cual participaron los usuarios previstos seleccionados), y que finalmente, fueron puestas a consideración del Comité de Seguimiento, cuatro fueron implementadas, seis no han sido implementadas o están descartadas y dos tienen evidencia de ser justificaciones de decisiones previas.

Finalmente, la medición acerca de la implementación de las recomendaciones permite establecer que la evaluación se desarrolló en un contexto en el cual se tenía interés por mejorar la política. Posterior a la evaluación se revelan cambios en la institucionalidad de la intervención como el CONPES 3943 de 2018 y la Resolución 2254 del 2017. Esto refleja que hubo apertura de los usuarios previstos hacia el cambio de la política con base en los resultados de la evaluación.

11. Análisis comparado

Este capítulo está destinado al abordaje de tres de los objetivos específicos de la presente investigación: brindar una estimación acerca del grado de efectividad actual que tiene la política pública de evaluación de políticas públicas; analizar la importancia relativa de los diversos estándares de calidad para la generación de efectividad por parte de los procesos de evaluación de política pública; y analizar la importancia relativa que tienen los componentes de la coordinación entre el proceso de evaluación con el de la política pública evaluada para la generación de efectividad por parte de los procesos de evaluación de política pública.

El capítulo comienza describiendo a grandes rasgos lo hallazgos relacionados con las tres variables de la hipótesis de investigación. Se inicia con una valoración general de la efectividad de la política de evaluación de políticas, la cual, como se podrá observar, está en consonancia con las conclusiones acerca de la efectividad de la política de evaluación que se obtuvieron con el análisis descriptivo del caso general correspondiente al segundo capítulo. Posteriormente se realiza una presentación general de los hallazgos relativos al cumplimiento de los estándares de calidad, y del grado de coordinación evidenciado en los casos seleccionados a comparar.

La última sección de este capítulo está destinada a la presentación de los resultados de los análisis comparados de similitud y de diferencia. Entrando en una lógica de inferencia explicativa, se abordará en estos análisis los objetivos relativos a la identificación de la importancia relativa que tienen los dos grupos de variables con las que se operacionalizaron las variables independientes de la hipótesis de investigación. Gracias a esta identificación, se podrá obtener el mecanismo causal que mejor explica la variabilidad de la efectividad de los procesos de evaluación, a partir del cual se podrá determinar la teorización actual más acertada al respecto.

11.1 Valoración de la efectividad de la política pública de evaluación de políticas

A un nivel descriptivo, teniendo en consideración que la operacionalización del concepto de efectividad, remite al uso de la evaluación (y este implica observar tanto un cambio en el modelo mental de los usuarios previstos, como un cambio institucional en la política evaluada que refleje las elecciones de estos), y que esto está en coherencia con el objetivo principal de la política pública, el cual es el mejoramiento de las intervenciones públicas (Decreto 1290 de 2014, art. 21) a partir de generar evidencia (art. 22) que sirva de insumo para la toma de decisiones (Decreto 1082 de 2015, Art. 2.2.7.3.1.4.), la valoración acerca de la efectividad de la política no es positiva.

Si bien es cierto que no es posible transpolar la valoración baja acerca del grado de efectividad de los procesos de evaluación que se obtuvo en esta investigación (de 101 recomendaciones estudiadas -83 de casos particulares y 18 de las evaluaciones de SINERGIA- solo 34 fueron implementadas, 26 si se consideran solo los casos seleccionados en el análisis comparado) a todos los procesos de evaluación realizados por SINERGIA hasta la fecha, a razón de que esta muestra no respondió a criterios de aleatoriedad, sino de muestreo intencional, si se puede decir que esta conclusión toma fuerza en tanto que tiene congruencia con las percepciones negativas acerca de la efectividad de la política que fueron presentadas en el caso general en el capítulo 2.

Dejando de lado las evaluaciones a SINERGIA[68] el mejor caso que evidencia efectividad de la Política de Evaluación de Políticas es la evaluación al Programa de Planes Departamentales de Agua en donde se implementaron 11 de 17 recomendaciones. Este caso es seguido por la evaluación al Programa de Reparación Colectiva que obtuvo 6 de 11 recomendaciones implementadas, y por la evaluación a la Política de Prevención y Control a la Contaminación del Aire con 4 de 12 recomendaciones implementadas. El resto de los casos no reflejan ni siquiera un grado que podría considerarse como de media efectividad. La evaluación a Conexiones Intradomiciliarias implementó 2 de 10; Política Nacional de Seguridad Alimentaria y Nutricional implementó 2 de 13; Medidas de Defensa Comercial implementó 1 de 9; Política de Retornos y Reubicaciones implementó 0 de 3; y Política de Atención Humanitaria 0 de 8. De ocho casos seleccionados, dos muestran un grado medio de efectividad, uno un grado medio-bajo, y cinco un grado claramente bajo de efectividad. Ningún caso puede decirse que refleje un alto grado

[68] Esto a razón de que estos procesos de evaluación no se usaron para el análisis comparado debido a que fueron realizadas hace un periodo considerable de tiempo, lo que aumenta el número de variables de control a utilizar para hacer una comparación coherente, con una institucionalidad considerablemente diferente.

de efectividad. Si se toma en cuenta las evaluaciones realizadas a SINERGIA, la del 2010 tendría un grado positivo de efectividad con 4 de 6 recomendaciones implementadas. La del 2013, en contraste, sólo tiene 4 de 12.

Adicional a esto, se encontró evidencia en 17 de las 83 recomendaciones del análisis comparado de corresponder a justificaciones de decisiones previas. Asimismo, se encontró que no solo las recomendaciones pueden servir como justificación de decisiones previas. El caso de la evaluación a la Política de Atención Humanitaria, por ejemplo, solo reflejó 2 de las 8 recomendaciones como justificaciones de decisiones previas, sin embargo, la finalidad de la evaluación orientada a validar el diseño de modificación, los comentarios del Comité de Seguimiento, y las entrevistas evidenciaron que los hallazgos de la evaluación, en general, fueron utilizados para justificar y legitimar un modelo de política con el que ya contaban los usuarios previstos.

11.2 Valoración del grado de calidad de los procesos de evaluación

Los procesos de evaluación estudiados muestran un alto grado de cumplimiento de los estándares correspondientes a reglas de alcance: fundamentación y finalidad de la evaluación; objeto y alcance; objetivos específicos; oportunidad, pertinencia y utilización; y desarrollo de capacidad. Únicamente se presentaron algunas falencias en el cumplimiento del estándar de desarrollo de capacidad en tres casos, y en el de oportunidad, pertinencia y utilización en un caso. Aun así, estas son falencias menores, en general, los procesos de evaluación satisfacen estos estándares desde su etapa de diseño, e institucionalmente esto es requerido de manera explícita.

En cuanto a la satisfacción de los estándares correspondientes a reglas de frontera, la situación es muy dividida. Por una parte, todos los procesos de evaluación llevan a cabo procesos de convocatoria transparentes y tienen requisitos exigentes frente al equipo evaluador a selección. Así mismo se garantiza una suficiencia de recursos en todos los procesos de evaluación. Los estándares de enfoque asociativo, consideración de evaluación conjunta y participación de los actores relevantes, tienen todos en común el requerimiento de incluir a los actores relevantes de la intervención en el proceso de evaluación.

Únicamente en la evaluación a la Política de Atención Humanitaria se incluyó a todos los actores relevantes, en el resto de los procesos, a pesar de que se incluyeron a las entidades que pueden ser consideradas como líderes de la formulación e implementación de las intervenciones, constantemente se observaron los límites de no incluir a todos los actores que tienen competencias en este proceso (por lo menos los del nivel nacional). Todas las 26 recomendaciones implementadas tuvieron como responsables a alguna entidad que fue incluida en el proceso de evaluación.

Adicionalmente, 42 de las 83 recomendaciones que no fueron implementadas, alguno de sus responsables (según el PTIR correspondiente) dejó de ser incluido

como usuario previsto en el proceso de evaluación, esta información se puede verificar en la siguiente tabla, la cual agrupa como filas a cada una de las recomendaciones estudiadas en la presente investigación, y ayuda a contrastar el grado de implementación de cada una, con la verificación de si los responsables que fueron asignados en el PTIR realmente fueron incluidos como usuarios previstos en el proceso de evaluación:

Tabla: Relación entre implementación de recomendación con los usuarios previstos involucrados

Recomendaciones	Responsables según PTIR	Usuarios involucrados	Implementación
Planes Departamentales de Agua			
1. Fortalecer la institucionalidad enmarcada en la Ley 142 de 1994 para apoyar a los municipios y prestadores, e intervenir donde sea pertinente para asegurar la prestación eficiente de los servicios	MVCT	MVCT, DNP	No imp.
2. Reglamentar el artículo 6º de la ley 142 referido al cumplimiento de las condiciones para que proceda la prestación directa de servicios por parte de los municipios	SSPD, CRA	MVCT, DNP	No imp.
3. Definir con mayor precisión el alcance del rol que desempeña el Gestor como ejecutor de los PDAs	MVCT	MVCT, DNP	Imp.
4. Definir el alcance del rol que desempeña el Gestor en temas ambientales.	MVCT	MVCT, DNP	Imp.
5. Definir estructuras tipo para los Gestores	MVCT	MVCT, DNP	Imp.
6. Promover contratos temporales con operadores especializados para brindar asistencia técnico-operativa y comercial	MVCT, gestores	MVCT, DNP	Imp.
7. Apoyar la planeación de largo plazo y articulada (Nación Departamentos Municipio)	MVCT, gestores, alcaldías	MVCT, DNP	Imp.
8. Formular mapeo de criterios para la distribución de los recursos de la Nación y los departamentos en los municipios vinculados al PAP-PDA de manera articulada (Nación, Departamentos, Municipio)	MVCT, gestores	MVCT, DNP	No imp.
9. Ejecutar recursos de la Nación vía PDAs y condicionarla al compromiso de recursos de los municipios	MVCT, alcaldías	MVCT, DNP	No imp.
10. Instituir un mecanismo de acceso a crédito para todos los municipios	MVCT, Gestor, Alcaldías	MVCT, DNP	No imp.
11. Implementar mecanismos de seguimiento periódico a los instrumentos de planeación	MVCT, Gestor, Alcaldías	MVCT, DNP	imp.

12. Fortalecer los procesos ejecutados por el mecanismo de viabilización de proyectos	MVCT MHCP	MVCT, DNP	Imp.
13. Articular la ejecución de inversiones a nivel regional	MVCT, alcaldías	MVCT, DNP	Imp.
14. Fortalecer las actividades de fortalecimiento institucional realizadas por el Gestor.	Gestores	MVCT, DNP	Imp.
15. Promover la sostenibilidad financiera de los prestadores de los servicios	Gestores	MVCT, DNP	No imp.
16. Establecer los alcances de los actores en las actividades de asistencia técnica.	MVCT, gestores	MVCT, DNP	Imp.
17. Redefinir el rol del Gestor frente al proceso de certificación de municipios	MVCT, gestores	MVCT, DNP	Imp.
Programa de Reparación Colectiva			
1. Revisar y redefinir los componentes del programa a través de la elaboración de un CONPES, un Decreto de Reparación Colectiva y un manual operativo.	No asignado. Competencia de Subcomité de Reparación Colectiva	UARIV, DPS, DNP.	Imp.
2. Desarrollar la dimensión colectiva de cada una de las medidas de reparación	No asignado. Competencia de la UARIV	UARIV, DPS, DNP.	Imp.
3. Clarificar los límites y las relaciones complementarias de la reparación con las medidas de desarrollo	No asignado. Competencia de la UARIV	UARIV, DPS, DNP.	Imp.
4. Elaborar un documento CONPES para el PRC	DNP, UARIV, DPS, MHCP, y Subcomité de Reparación Colectiva	UARIV, DPS, DNP.	No imp.
5. Elaboración del Decreto de Reparación Colectiva	DNP, UARIV, Min. Interior, Min. Justicia, y Subcomité de Reparación Colectiva	UARIV, DPS, DNP.	Imp.
6. Elaborar el manual operativo del PRC	UARIV, DNP, y DPS	UARIV, DPS, DNP.	Imp.
7. Definir una batería de indicadores de resultados	UARIV, DNP, y DPS	UARIV, DPS, DNP.	No imp.

8. Ajustar la composición del Subcomité de Reparación Colectiva	Comité Ejecutivo del SNARIV	UARIV, DPS, DNP.	No imp.
9. Alinear las metas y los objetivos del subcomité con las metas y los objetivos del programa a nivel de resultados	UARIV, DPS y Subcomité de Reparación Colectiva	UARIV, DPS, DNP.	No imp.
10. Ajustar la participación de las entidades del SNARIV en la ruta de reparación colectiva	UARIV y Subcomité de Reparación Colectiva	UARIV, DPS, DNP.	No imp.
11. Estimar el esfuerzo fiscal que debe realizar el Estado para implementar la RC	UARIV, DNP, MHCP y DPS	UARIV, DPS, DNP.	Imp.
Política de Prevención y Control de la Contaminación del Aire			
1. Fortalecimiento de la coordinación interinstitucional para mejorar la efectividad de la política.	MADS, Min. Minas, Min. Transporte y Min. Salud	MADS, DNP.	JDP
2. Incluir en la PPCCA lineamientos para diferenciar la implementación de las medidas de prevención y control según características geográficas y nivel de incumplimiento de la norma nacional.	MADS, Min. Salud	MADS, DNP.	No imp.
3. Priorizar medidas de política en contaminantes críticos y establecer normas de calidad del aire resultado de análisis costo-beneficio	MADS	MADS, DNP.	JDP
4. Socialización masiva regional y nacional de estado de calidad del aire, costos de salud, metas de programas de reducción y avances de los gobiernos en mejoramiento de calidad del aire	MADS, autoridades regionales ambientales	MADS, DNP.	Imp.
5. Ampliar y modificar la regulación de fuentes fijas basada en permiso de emisiones	MADS, MinCIT y autoridades regionales ambientales	MADS, DNP.	No imp.
6. Reglamentar instrumentos económicos para alcanzar metas de reducción de emisiones en áreas donde se incumple la norma diaria y donde las fuentes fijas contribuyan significativamente a la problemática.	MADS	MADS, DNP.	imp.
7. Incentivos tributarios para fuentes fijas alineados (simplificados) con programas de reconversión tecnológica	MADS, ANLA	MADS, DNP.	No imp.

8. Racionalizar obligaciones de monitoreo de calidad del aire y emisiones a los responsables de fuentes fijas	MADS, IDEAM	MADS, DNP.	No imp.
9. Etiquetado de vehículos (particulares, pasajeros y carga) para permitir alinear restricciones de movilidad con nivel de emisión vehicular	MADS y Min. Transporte	MADS, DNP.	No imp.
10. Impuestos de vehículos basados en edad y tecnología, para incentivar renovación flota	Congreso de la República, MHCP	MADS, DNP.	No imp.
11. Inclusión de análisis costo-beneficio ambiental en decisiones de renovación y adopción tecnologías vehiculares de transporte público masivo y transporte público integrado en áreas urbanas.	Min. Transporte, MADS	MADS, DNP.	Imp.
12. Establecer calendarios y metas de renovación del parque automotor diesel como pre-requisito al suministro de diesel de mejor calidad.	Min. Minas, MADS, Min. Transporte.	MADS, DNP.	Imp.
Programa de Conexiones Intradomiciliarias			
1. Seguimiento indicadores de salud	MVCT, Min. Salud	MVCT, FONADE, DNP.	No imp.
2. Verificar el requisito de calidad del servicio	MVCT	MVCT, FONADE, DNP.	JDP
3. Adecuación estética de los baños	MVCT	MVCT, FONADE, DNP.	No imp.
4. Capacitar a la población de los barrios beneficiaros en oficios demandados por el programa	MVCT, SENA, FONADE	MVCT, FONADE, DNP.	JDP
5. Exigir que el contratista entre en consorcio con un proveedor de insumos	FONADE, contratista	MVCT, FONADE, DNP.	No imp.
6. Realizar una encuesta de seguimiento a la población beneficiaria	MVCT, FONADE	MVCT, FONADE, DNP.	JDP
7. Inclusión de DSR en las capacitaciones y empoderamiento de la mujer	MVCT	MVCT, FONADE, DNP.	No imp.
8. Facilitar y flexibilizar las formas de pago del servicio de agua	MVCT, ESP	MVCT, FONADE, DNP.	Imp.

9. Seguimiento y control a las socializaciones y capacitaciones realizadas por las ESP y las Alcaldías Municipales	MVCT	MVCT, FONADE, DNP.	imp.
10. Interacción entre actores y sostenibilidad del programa	MVCT, ESP	MVCT, FONADE, DNP.	JDP
Política Nacional de Seguridad Alimentaria y Nutricional			
1. Priorizar la agenda intersectorial en el trabajo de la CISAN	Presidencia y Secretaría Técnica de la CISAN	DPS, MADR, MSPS, DNP	No imp.
2. Interpretar los lineamientos SAN de la Corte Constitucional en el Plan de Acción	Presidencia y Secretaría Técnica de la CISAN	DPS, MADR, MSPS, DNP	Imp.
3. Construir e implementar las cadenas de valor inter-sectoriales	Entidades CISAN	DPS, MADR, MSPS, DNP	No imp.
4. Agendar preferentemente temas intersectoriales en el temario de la CISAN	Presidencia y Secretaría Técnica de la CISAN	DPS, MADR, MSPS, DNP	Imp.
5. Asegurar el engranaje de las intervenciones sectoriales con los planes de acción anuales de las entidades nacionales	Entidades CISAN	DPS, MADR, MSPS, DNP	No imp.
6. Impulsar el desarrollo de enfoque “De la granja a la mesa”	MADR, MSPS, MADS, INVIMA, ICA	DPS, MADR, MSPS, DNP	JDP
7. Fortalecer el seguimiento de la política	Secretaría técnica CISAN y OSAN	DPS, MADR, MSPS, DNP	No imp.
8. Precisar las entidades y dependencias responsables de la implementación de las intervenciones del PNSAN que no tienen responsable	Secretaría técnica en coordinación con las entidades de la CISAN	DPS, MADR, MSPS, DNP	No imp.
9. Recoger las prioridades de la SAN definidas por los municipios y enfatizar las prioridades nacionales que no son prioritarias para los municipios.	Entidades CISAN	DPS, MADR, MSPS, DNP	No imp.

10. Convocar a la CISAN para definir una respuesta frente al comportamiento de las metas críticas evidenciadas en esta evaluación	Entidades CISAN	DPS, MADR, MSPS, DNP	JDP
11. Una presidencia inter-sectorial para la CISAN.	Entidades CISAN	DPS, MADR, MSPS, DNP	JDP
12. Fortalecer el OSAN	Secretaría Técnica	DPS, MADR, MSPS, DNP	JDP
13. Integrantes CISAN (mesa técnica precios de alimentos)	Secretaría Técnica	DPS, MADR, MSPS, DNP	No imp.
Medidas de Defensa Comercial			
1. Mejorar procedimientos relativos a la solicitud	MinCIT	MinCIT, DNP.	JDP
2. Paquete de mejoras que permitan destrabar el cuello de botella en la etapa de investigación	MinCIT	MinCIT, DNP.	JDP
3. Mejorar en la ejecución del procedimiento de recomendación	MinCIT	MinCIT, DNP.	JDP
4. Instaurar ajustes al procedimiento de seguimiento de las medidas de defensa comercial	MinCIT	MinCIT, DNP.	JDP
5. Establecer unos lineamientos de manejo de confidencialidad que sean iguales para todas las partes	MinCIT	MinCIT, DNP.	No imp.
6. Elevar el rango de la SPC a directivo para que pueda dar un mejor manejo a las presiones externas	Gobierno Nacional	MinCIT, DNP.	No imp.
7. Limitar los procesos que serán transmitidos a instancias de recomendación y decisión, donde sólo se discutan aquellos casos que correspondan a situaciones que afecten el interés nacional	MinCIT	MinCIT, DNP.	No imp.
8. Reglamentar que dentro de los compromisos trazados por el solicitante cuando se le otorga una medida de salvaguardia, la SPC tenga derecho a acceder a toda la información que considere necesaria para realizar sus investigaciones	MincIT	MinCIT, DNP.	JDP
9. Mejorar la dinámica de la recomendación y decisión en la operación, donde se presenta la apertura de espacio a partes involucradas en instancias de recomendación y decisión, la supresión de las actividades relacionadas con defensa comercial por parte del CSCE y la creación de un reglamento interno para las instancias encargadas de esta tarea.	MinCIT	MinCIT, DNP.	Imp
Política de Retornos y Reubicaciones			

1. Redefinición conceptual del proceso de acompañamiento estatal.	UARIV, Min. Interior, Presidencia, URT, MHCP, DAFP	UARIV, DPS, DNP.	No imp.
2. Incentivos para el involucramiento local en soluciones duraderas.	UARIV, DPS, DNP, Min. Interior, MHCP, Presidencia, DAFP	UARIV, DPS, DNP.	JDP
3. Reingeniería de los flujos de conocimiento.	UARIV, DPS, Min. Interior, DNP, Min. Defensa, Defensoría, Procuraduría.	UARIV, DPS, DNP.	No imp.
Política de Atención Humanitaria			
1. Mejoramiento de la calidad en la información de registro	UARIV, Ministerio Público	UARIV, DNP.	JDP
2. Verificación in situ de la caracterización	UARIV	UARIV, DNP.	No imp.
3. Establecimiento de un procedimiento progresivo de saluda.	UARIV	UARIV, DNP.	No imp.
4. Cubrimiento del riesgo discal por decisiones judiciales	UARIV	UARIV, DNP.	No imp.
5. Financiamiento de corto plazo para garantizar las fechas de pago	UARIV	UARIV, DNP.	No imp.
6. Desarrollo del programa de pedagogía del pago	UARIV	UARIV, DNP.	No imp.
7. Desarrollar el proceso de remisión y comunicación a la víctima de dicha remisión a las demás medidas de atención, asistencia y reparación	UARIV	UARIV, DNP.	JDP
8. Optimización de procesos y procedimientos	UARIV	UARIV, DNP.	No imp.

Se utiliza “imp.” para calificar la recomendación como implementada; “No imp.” Como no implementada, y “JDP” para señalarla como justificación de decisión previa.

En relación con las reglas de elección y su cumplimiento, en la gran mayoría de casos se observa un adecuado cumplimiento de los estándares que giran alrededor de la calidad que deben tener los productos de las evaluaciones. Por lo general, la labor del Comité de Diseño y del Comité de Seguimiento es muy eficiente para requerir productos de buena calidad. Esto se ve reflejado de manera particular en

la naturaleza de los comentarios realizados por el Comité de Seguimiento y que quedaron plasmados en los conceptos técnicos que elabora la DSEPP de los productos de la evaluación, en estos comentarios se observa un elevado grado de exigencia técnica que se orienta principalmente hacia el cumplimiento de los requerimientos que se consignaron en el pliego de condiciones y en el contrato. Sin embargo, hay algunas excepciones. El estándar de menciones de discrepancias entre miembros del equipo evaluador rara vez se cumple, y por lo general, los procesos de evaluación fallan en uno o dos de los 16 estándares que corresponden a reglas de elección.

El mejor desempeño en cuanto al cumplimiento de estándares se da en las reglas de control. En todos los procesos de evaluación estudiados se observa un cumplimiento de todos los componentes de los estándares de control de calidad e inclusión de observaciones de los implicados. Se observa un amplio reconocimiento por parte de los usuarios previstos de la labor que ejerce la DSEPP en ese sentido. Sin embargo, la combinación de la competencia del Comité de Seguimiento en cuanto a la supervisión técnica de los productos, con las cláusulas que funcionan como reglas de control en el contrato (mecanismos de cobertura del riesgo, cláusula penal compensatoria y multas), con el corto plazo que tiene el consultor para entregar los productos corregidos, y con el hecho de que se deban corregir los productos antes del pago, genera que el consultor tenga fuertes incentivos institucionales para incluir comentarios que van en contravía de los estándares correspondiente a reglas de posición (independencia del evaluador y proceso evaluativo libre y abierto). Como se observó, existen muchos casos en donde se requirió hacer recomendaciones o conclusiones en un sentido particular, lo que el consultor prefirió incluir antes de generar una discusión que retrasaría su pago.

En contraste con el perfecto cumplimiento de los estándares correspondientes a las reglas de control, únicamente la mitad de estos procesos cumplieron con los estándares categorizados como reglas de posición. Finalmente, en términos de reglas de información, también se presenta un muy buen desempeño. Por lo general, el pliego de condiciones definitivo y el contrato son herramientas de muy buena calidad producidas por el Comité de Diseño de la evaluación. También se observa un muy buen grado de difusión de las evaluaciones, tanto externamente con el portal de SINERGIA, como internamente con la relación continua que se da entre evaluadores y usuarios previstos en el Comité de Seguimiento, y con la difusión que se les realiza a los directivos de las entidades que lo conforman. Solo se presentan unas fallas menores en dos casos particulares relacionadas con problemas de claridad del informe, lo que restringió su utilidad.

11.3 Valoración del grado de coordinación como patrones de interacción requeridos en los procesos de evaluación realizados por la DSEPP

En cuanto a la coordinación/adaptación mutua que se definió como segunda variable independiente, y que busca poner a prueba los postulados de Patton para explicar la variación de la efectividad de los procesos de evaluación, en un nivel descriptivo se puede establecer que por lo general el DNP y los consultores seleccionados como evaluadores son reactivos frente a las preocupaciones e inquietudes de los usuarios previstos en temas como la identificación de las necesidades de información y selección de preguntas orientadoras; la definición conjunta de los usos de las evaluaciones; los distintos debates metodológicos para la selección del enfoque de la evaluación; las distintas adaptaciones en el método que se dan a lo largo del proceso de evaluación (sólo en las evaluaciones de Conexiones Intradomiciliarias y en Atención Humanitaria se presentaron problemas en la reactividad del evaluador frente a estos temas).

Este mismo grado de reactividad y adaptabilidad no se ve reflejado en el involucramiento de los usuarios previstos en la interpretación de los hallazgos; en el involucramiento de los usuarios en la construcción de recomendaciones; en la construcción conjunta de un seguimiento a la implementación de los hallazgos de la evaluación; o en la búsqueda proactiva del uso de la evaluación que debería estar a cargo del evaluador o del DNP (teniendo en consideración que la propiedad de los productos es de este último según los contratos de consultoría). Tres casos tuvieron problemas para involucrar a usuarios en la interpretación de los hallazgos y cuatro en el involucramiento en la construcción de recomendaciones.

Las mediciones más críticas se presentan en la ausencia total de un seguimiento sistemático a la implementación de los hallazgos de la evaluación, en cuyo mejor caso se limitó a la formulación de un conjunto de indicadores que nunca se utilizaron. Adicionalmente, solo un caso presentó cierta evidencia de una búsqueda proactiva del uso por parte del evaluador, pero por lo general, la ruptura de la relación contractual entre el equipo evaluador y la DSEPP, y la ausencia de una comunicación entre esta última y los usuarios previstos luego de la difusión interna de los hallazgos de la evaluación, restringen fuertemente la posibilidad de que esta búsqueda sirva a la coordinación y adaptación mutua entre los participantes de las arenas de acción analizadas.

Si bien la medición más crítica se presenta en estas dos dimensiones, el mayor impacto negativo en términos de coordinación se genera por fallas en el hallazgo e involucramiento como usuarios previstos de todos los actores que realmente son necesarios para el uso de la evaluación. Más allá del cumplimiento de los requerimientos de los estándares que corresponden a reglas de frontera (los cuales no tienen ningún tipo de impacto en la variable dependiente, tal como se evidencia

en la siguiente sección), la ausencia de asignación del rol de usuario previsto en el Comité de Seguimiento a muchos actores de las intervenciones, que fueron relegados a un papel consultivo, fue la razón del surgimiento de conclusiones y recomendaciones irrelevantes. Esta conclusión toma fuerza al observar que, en los procesos de evaluación de las intervenciones de naturaleza intersectorial, aquellas recomendaciones en las que se asignó como responsable a entidades que no fueron incluidas en el proceso de evaluación como usuarios previstos, casi su totalidad dejaron de ser implementadas.

Una salvedad que debe hacerse es que en ningún caso el involucramiento de los usuarios en la interpretación de los hallazgos o en la construcción de recomendaciones debería implicar un impacto negativo en los estándares de independencia del evaluador o de proceso evaluativo libre y abierto, puesto que este involucramiento puede estar (y de hecho en la mayoría de casos está) restringido al requerimiento de la consideración de información extra, el aumento de rigurosidad, o el fortalecimiento con evidencia de argumentos.

11.4 Análisis comparado con lógica de similitud y diferencia

En la siguiente tabla sintetiza todas las mediciones que se realizaron para el análisis comparado, cada fila corresponde a un caso seleccionado. En las primeras columnas, se muestra las mediciones relacionadas con la variable dependiente, se continúa con la presentación de la medición de los estándares de calidad, en donde se establecen tres tipos de medición: “S” para designar una satisfacción completa de todos los componentes del estándar; “PS” para denotar una parcial satisfacción del estándar; y “NS” para asignar la no satisfacción del estándar. En la parte relacionada con las mediciones acerca de la coordinación, existen tres calificaciones también: “A” para calificar como alta la presencia de una coordinación como adaptación mutua en ese particular tema de la evaluación; “M” para designar una presencia media; y “B” para designar una baja presencia.

Tal como puede observarse en la tabla, si nos restringimos a un análisis comparado de similitud, en donde la lógica de interpretación gira alrededor de la idea de que, si una condición X es necesaria para que exista Y, entonces todos los casos en los que exista Y deben ofrecer también la presencia de X, se deben entonces observar ¿cuáles son los factores comunes que tienen los tres casos en donde existe efectividad de los procesos de evaluación? Tal como se puede observar, sólo los casos de Planes Departamentales de Agua, Reparación y Política de Prevención y Control de la Contaminación del Aire pueden considerarse como relativamente efectivos en cuanto al uso de la evaluación. En este sentido, las variables que adquieren fuerza como explicación de la efectividad de los procesos de evaluación son estas:

1. Fundamentación y finalidad de la evaluación
2. Objeto y alcance de la evaluación
3. Objetivos Específicos

4. Oportunidad, pertinencia y utilización
5. Equipo de evaluación
6. Recursos
7. Coordinación y alineamiento
8. Selección de enfoque y método
9. Selección y aplicación de criterios de evaluación
10. Cuestiones a responder
11. Claridad del análisis
12. Contexto de la intervención
13. Validez y fiabilidad de fuentes de información
14. Consulta y protección de implicados
15. Control de calidad
16. Inclusión de observaciones de los implicados
17. Documentación sobre finalidad y expectativa
18. Hallar e involucrar a usuarios en proceso de evaluación
19. Escuchar preocupaciones y demandas de usuarios
20. Revisar opciones de preguntas de evaluación
21. Trabajo conjunto sobre usos
22. Discutir debates metodológicos
23. Adaptación del diseño metodológico
24. Involucramiento de usuarios en interpretación de hallazgos
25. Involucramiento de usuarios en la construcción de recomendaciones

A pesar de que el análisis de similitud nos puede hablar de los factores que pueden considerarse como necesarios para que se utilicen los resultados de las evaluaciones en la toma de decisiones, este tipo de análisis no tiene la suficiente fuerza como para poder generar inferencias alrededor de los factores que determinan la efectividad de la evaluación. En particular, el análisis de similitud no nos dice nada de aquellos casos en donde se encuentran las variables anteriormente mencionadas, pero en las que no se da el resultado esperado. Por esta razón, es necesario avanzar hacia un análisis comparado más profundo.

Tabla comparativa

	Efectividad		Reglas de alcance y resultados potenciales				
	Implementación de recomendaciones	Justificación de decisión previa	Fundamentación y finalidad de la evaluación	Objeto y alcance de la evaluación	Objetivos Específicos	Oportunidad, pertinencia y utilización	Desarrollo de capacidad
Planes Departamentales de Agua	11 de 17	0 de 17	S	S	S	S	S
Reparación Colectiva	6 de 11	0 de 11	S	S	S	S	S
Contaminación del aire	4 de 12	2 de 12	S	S	S	S	PS
Conexiones intradomiciliarias	2 de 10	4 de 10	S	S	S	PS	PS
Seguridad Alimentaria y Nutricional	2 de 13	4 de 13	S	S	S	S	PS
Medidas de Defensa Comercial	1 de 9	5 de 9	S	S	S	S	S
Retornos y Reubicaciones	0 de 3	1 de 3	S	S	S	S	S
Atención humanitaria	0 de 8	2 de 8	S	S	S	S	S

	Reglas de frontera y participantes					Reglas de elección y acciones de los participantes					
	Participación de actores relevantes	Consideración de evaluación conjunta	Enfoque asociativo	Equipo de evaluación	Recursos	Evaluabilidad	Coordinación y alineamiento	Selección de enfoque y método	Selección y aplicación de criterios de evaluación	Cuestiones a responder	Respuestas a preguntas evaluativas
Planes Departamentales de Agua	NS	NS	NS	S	S	NS	S	S	S	S	PS
Reparación Colectiva	NS	NS	NS	S	S	PS	S	S	S	S	S
Contaminación del aire	PS	PS	PS	S	S	S	S	S	S	S	S
Conexiones intradomiciliarias	NS	NS	NS	S	S	S	PS	S	S	S	S
Seguridad Alimentaria y Nutricional	PS	PS	PS	S	S	S	S	S	S	S	S
Medidas de Defensa Comercial	PS	PS	PS	S	S	S	S	S	S	S	S
Retornos y Reubicaciones	PS	PS	PS	S	S	S	S	S	S	S	S
Atención humanitaria	S	S	S	S	S	S	S	S	S	S	PS

	Reglas de elección y acciones de los participantes									
	Claridad del análisis	Claridad y representatividad del resumen	Contexto de la intervención	Lógica de la intervención	Explicación de la metodología	Validez y fiabilidad de fuentes de información	Mención de cambios y límites de metodología	Consulta y protección de implicados	Mención de discrepancias entre evaluadores	Evaluación dentro del plazo y presupuesto
Planes Departamentales de Agua	S	NS	S	S	S	S	PS	S	NS	S
Reparación Colectiva	S	S	S	S	S	S	S	S	S	NS
Contaminación del aire	S	S	S	NS	PS	S	S	S	NS	S
Conexiones intradomiciliarias	S	S	S	S	S	S	S	S	NS	S
Seguridad Alimentaria y Nutricional	S	S	S	S	PS	S	S	S	NS	S
Medidas de Defensa Comercial	PS	S	S	S	S	S	NS	S	S	S
Retornos y Reubicaciones	S	S	S	PS	S	S	S	S	NS	NS
Atención humanitaria	S	S	S	S	S	S	S	S	NS	S

	Reglas de agregación		Reglas de posición		Reglas de información		
	Control de calidad	Inclusión de observaciones de los implicados	Independencia de los evaluadores	Proceso evaluativo libre y abierto	Documentación sobre finalidad y expectativa	Informe de la evaluación	Difusión de la evaluación
Planes Departamentales de Agua	S	S	NS	NS	S	S	PS
Reparación Colectiva	S	S	S	S	S	S	S
Contaminación del aire	S	S	NS	NS	S	NS	S
Conexiones intradomiciliarias	S	S	S	S	S	S	S
Seguridad Alimentaria y Nutricional	S	S	NS	NS	S	S	S
Medidas de Defensa Comercial	S	S	S	S	S	S	S
Retornos y Reubicaciones	S	S	S	S	S	S	S
Atención humanitaria	S	S	PS	PS	S	PS	S

En esta sección de la tabla comparativa se utiliza: “S” para designar una satisfacción completa de todos los componentes del estándar; “PS” para denotar una parcial satisfacción del estándar; y “NS” para asignar la no satisfacción del estándar.

	Coordinación/adaptación mutua							
	Recomendaciones implementadas a las que se involucró a alguno de su(s) responsable(s)	Recomendaciones a las que no se involucró alguno de sus responsables	Hallar e involucrar a usuarios en proceso de ev.	Escuchar preocupaciones y demandas de usuarios (reactividad)	Revisar opciones de preguntas de ev.	Trabajo conjunto sobre usos	Discutir debates metodológicos	Adaptación del diseño metodológico
Planes Departamentales de Agua	11 de 11	7 de 17	A	A	A	A	A	A
Reparación Colectiva	6 de 6	9 de 11	A	A	A	A	A	A
Contaminación del aire	4 de 4	10 de 12	A	A	A	A	A	A
Conexiones intradomiciliarias	2 de 2	4 de 10	M	B	A	A	M	B
Seguridad Alimentaria y Nutricional	2 de 2	7 de 13	M	A	A	A	A	A
Medidas de Defensa Comercial	1 de 1	1 de 9	A	M	A	M	A	A
Retornos y Reubicaciones	0 de 0	3 de 3	B	A	A	A	A	A
Atención humanitaria	0 de 0	1 de 8	A	A	A	A	A	M

	Coordinación/adaptación mutua				Nueva variable de control	
	Involucrar usuarios en interpretación de hallazgos	Involucrar usuarios en construcción de recomendación	Plan conjunto de seguimiento	Búsqueda proactiva de uso	Apertura de funcionarios a cargo de la política hacia el cambio de la institucionalidad de la intervención	
Planes Departamentales de Agua	A	A	B	B	A	Contexto con fuerte interés en la reformulación del programa a través de un decreto (a 2019 solo se cuenta con proyecto de decreto, pero con una alta probabilidad de ser decretado, a julio de 2019 está en proceso de sanción presidencial)
Reparación Colectiva	A	A	B	B	A	Elaboración de proyecto de decreto, resolución y manuales metodológicos, permiten establecer alto grado de oportunidad y relevancia
Contaminación del aire	A	A	B	B	A	Contexto con un fuerte interés de reformulación de la PPCCA a través de un CONPES (3943 de 2018)
Conexiones intradomiciliarias	B	B	M	B	A	Alto grado de apertura al cambio en la institucionalidad del programa, se formuló Guía Metodológica
Seguridad Alimentaria y Nutricional	A	B	B	B	B	Hubo apertura en nivel técnico, pero no en nivel directivo.
Medidas de Defensa Comercial	M	B	B	B	A	Varios cambios institucionales luego de la evaluación reflejan buen grado de apertura a cambios en la institucionalidad
Retornos y Reubicaciones	A	A	B	A	B	A julio de 2019 se requieren cambios en la Ley 1448 de 2011 antes de modificar la institucionalidad de la política en específico.
Atención humanitaria	A	A	B	B	B	Se expidió en un decreto el nuevo diseño institucional antes de tener los resultados de la evaluación, nuevas decisiones descartadas.

En la presente sección de la tabla comparativa, se utiliza “A” para calificar como alta la presencia de una coordinación como adaptación mutua en ese particular tema de la evaluación; “M” para designar una presencia media; y “B” para designar una baja presencia.

Otro problema del método de similitud es que a pesar de que genera evidencia a favor de la hipótesis que se pone a prueba en esta investigación, no nos permite discriminar entre las teorizaciones centradas en la calidad del proceso de evaluación y la teorización del U-FE de Patton, más centrada en el tipo específico de relacionamiento entre los participantes. En consecuencia, el método de similitud tampoco permitiría observar la relevancia relativa que tienen los componentes de esas teorizaciones.

Si avanzamos en la utilización del método comparado de diferencia, cuya lógica es que: "… todos los casos que presentan el resultado de interés deben también compartir una característica explicativa común (con suerte, aquella característica anticipada por nuestra hipótesis), mientras que los casos negativos deben carecer de este atributo" (Pérez, 2007, p. 10). Una aproximación a la comparación de la variabilidad entre la efectividad y la aplicación de los estándares profesionales de calidad refleja que no existe una sola variable dentro de este conjunto teórico que tenga una variabilidad igual a la de efectividad, esto significa que ninguna variable es independientemente suficiente para explicar el cambio en la variable de efectividad. Lo cual, también nos dice muy poco acerca de los factores que causan la efectividad, si la hipótesis es verdadera o no, o de cuál de las dos teorizaciones es mejor para explicar la efectividad.

En un nivel muy descriptivo, se debe resaltar (acerca de la importancia de los estándares acerca de la calidad de los productos) que el proceso de evaluación con mayor número de estándares insatisfechos clasificados como reglas de alcance, fue el proceso de evaluación con mayor grado de efectividad. Adicionalmente, los estándares de participación de actores relevantes, consideración de evaluación conjunta, enfoque asociativo, independencia del evaluador, y proceso evaluativo libre y abierto, también son variables que varían de una manera disímil a la efectividad de la evaluación, por lo que pierden relevancia como posibles explicaciones del resultado en cuestión.

En contraste, los componentes seleccionados del modelo Utilization-Focused Evaluation (U-FE) que permiten caracterizar los componentes de una coordinación como adaptación mutua en los procesos de evaluación evidencian un mayor grado de homogeneidad con relación a la variación de la efectividad. En una mirada preliminar del modo en el que varían las variables del U-FE con relación a la efectividad, parece que la conjunción entre el hallazgo e involucramiento de los usuarios previstos por parte del DNP y la reactividad del DNP y el equipo evaluador a las preocupaciones e inquietudes de los usuarios previstos en las etapas de: la definición del uso, las preguntas, la metodología, la adaptación de la metodología, interpretación de hallazgos, hasta la construcción de recomendaciones, parece ser la posible explicación de la variabilidad de la efectividad. Esto únicamente aplicando la lógica del método de diferencia a una posible configuración de variables causales.

Sin embargo, el caso correspondiente a la evaluación de la Política de Atención Humanitaria desafía esta explicación al tener una calificación casi perfecta en las variables del U-FE mencionadas, a la vez que presenta un resultado negativo. Esto

generó la siguiente pregunta: ¿qué diferencia existe entre los casos que presentan un resultado positivo y el proceso de evaluación de la Política de Atención Humanitaria? La mayor diferencia se encuentra en que de manera concomitante a la realización de la evaluación, se estaba diseñando la nueva institucionalidad de la Política de Atención Humanitaria a través del Decreto 2569 de 2014. De hecho, la finalización del diseño del decreto terminó cuatro meses antes de que se tuvieran disponibles los resultados de la evaluación.

Más allá de las implicaciones que esto tenga en cuanto a la oportunidad de la evaluación o a la utilización de las recomendaciones como justificaciones de decisiones previas, este fenómeno puede enmarcarse mejor en lo que Patton denomina la apertura de los usuarios de la evaluación hacia el cambio de la intervención (2012, p. 129) con base en los resultados de la evaluación, factor esencial para el mejoramiento de misma[69].

En términos de Ostrom, esto significa la apertura hacia el cambio institucional. Entonces, si se quiere que una evaluación se utilice debe existir por lo menos una situación de acción en la intervención evaluada, en donde se den las siguientes características: a) los participantes tienen como resultado potencial la reformulación (o el cambio) de la política pública evaluada; b) los participantes tienen como acción disponible la transformación de su modelo mental de la intervención gracias a los hallazgos de la evaluación; c) los participantes tienen una valoración positiva acerca de esta transformación, y d) Los participantes tienen como acción disponible la introducción de esas transformaciones dentro de la institucionalidad[70] de la política.

Esta nueva variable se agrega en las dos últimas columnas de la matriz comparativa, en donde además de la calificación correspondiente, se sintetiza las razones por las cuales se brindó dicha calificación al estado de la variable. Esta medición es consistente con la variabilidad de la efectividad puesto que, por un lado, los casos efectivos también presentan un alto grado de apertura hacia el cambio institucional, y por otro, varios de los casos con baja efectividad tienen también una baja medición de apertura al cambio.

---

[69] Todas las evaluaciones realizadas en el marco de SINERGIA por parte de la DSEPP y sus consultores son, por lo menos parcialmente, orientadas al mejoramiento de la intervención (también llamada evaluación de tipo formativo), lo cual se evidencia en que el último producto de cada proceso corresponda al Plan de Transferencia e Implementación de Recomendaciones.

[70] La información que proveen las arenas de acción de nivel operativo correspondientes a la ejecución de los procesos de evaluación, al tener una finalidad de mejoramiento o formativa, está orientada a incidir en la modificación de reglas, normas o estrategias (instituciones) de las situaciones de acción de la política evaluada. Las recomendaciones nunca se establecen como reglas en sí mismas, que determinan las acciones de los participantes de las situaciones de acción de nivel operativo en la política evaluada, es por esto por lo que las transformaciones del modelo mental de los usuarios siempre quedan plasmadas en cambios institucionales.

Se observa que las evaluaciones que tienen menor grado de efectividad son las que tienen mayor evidencia de orientar sus recomendaciones como justificaciones o modos de legitimación de decisiones previas. Únicamente en el caso de la evaluación de la Política de Prevención y Control de la Contaminación del aire se encontró evidencia de este tipo de recomendaciones, mientras que en los otros dos casos efectivos no hubo un solo caso de recomendación como justificación de decisión previa.

En este sentido, la conclusión preliminar que surge utilizando únicamente la lógica del análisis comparado de diferencia es que la conjunción entre el hallazgo e involucramiento de los usuarios previstos por parte del DNP, la reactividad del DNP y el equipo evaluador a las preocupaciones e inquietudes de los usuarios previstos en las distintas etapas del proceso de evaluación, más la apertura de los usuarios previstos al cambio institucional, y una ausencia de (o un rechazo a) utilización de las recomendaciones de la evaluación como modo de justificación o legitimación de decisiones previas, es una posible explicación de la variabilidad de la efectividad de los procesos de evaluación de políticas públicas, lo que dejaría por fuera a los estándares de calidad de la OCDE.

Finalmente, la última de las conclusiones preliminares que permite obtener la lógica del método comparado de diferencia, es que las variables de Patton de planeación conjunta del seguimiento a la implementación de los hallazgos de la evaluación y la búsqueda proactiva del uso por parte del evaluador (o el DNP en este caso), no surgen como variables causales plausibles, esto a razón de que en todos los casos positivos en términos de efectividad no se realizaron estas variables.

11.5 Análisis comparado con conjuntos nítidos

En términos estrictos, a pesar de que la lógica de similitud y de diferencia brindan importantes ideas preliminares acerca de las variables que toman individualmente relevancia como posibles explicaciones de la efectividad de los procesos de evaluación, también es cierto que no son herramientas que permitan establecer que un conjunto de variables (o una configuración causal particular) es suficiente para explicar la variación del resultado en cuestión. En este sentido, es necesario avanzar a una aplicación más rigurosa del método comparado, para lo cual se ha seleccionado el método QCA (Qualitative Comparative Analysis) con conjuntos nítidos desarrollado por Ragin.

El primer paso que hay que tener en cuenta para aplicar el método comparado con conjuntos nítidos es la realización de una apropiada selección de variables. Como puede apreciarse en la tabla de la sección anterior, en este punto de la investigación existen 45 variables independientes. Teniendo en cuenta que el método comparado

con conjuntos nítidos incluye la realización de valoración de consistencia por cada posible configuración, y que el número posible de configuraciones posibles teniendo 45 variables y 3 calificaciones posibles para cada uno es demasiado elevado (es 3^{45}), es necesario reducir el número de variables. Según Rihoux y Ragin (2009, pp. 25-28), los elementos que deben ser tenidos en cuenta para la selección de variables para la elaboración de un modelo comparativo son:

- La selección de variables debe estar guiada a partir de criterios teóricos, por lo que el investigador debe tratar de reducir su perspectiva a solo unas cuantas teorías “nucleares” o “centrales”.
- No incluir factores que no varían a lo largo de los casos seleccionados.
- Mantener el número de factores causales relativamente bajo (de 6 a 8 variables causales).
- Por cada condición es necesario tener hipótesis claras acerca de su conexión con el resultado.

En segundo lugar, para la aplicación de este método es necesario reducir el número de calificación posibles dentro de cada una de las variables (incluyendo la variable dependiente), para esto se utiliza las convenciones y operaciones del algebra booleana, en la cual solo deben existir dos calificaciones:

- Casos positivos en la medición de la variable son representados con “1”
- Casos en donde no exista la variable o esta presente una baja calificación son representados con “0”

Utilizar este sistema binario también permite realizar la minimización booleana. Esta minimización consiste en que, si dos expresiones booleanas difieren únicamente en una condición causal, pero aun así producen el mismo resultado, entonces la condición causal que la distingue puede ser considerada como irrelevante, por lo que puede ser removida para crear una expresión combinada más simple (p. 35). Este ejercicio es necesario para satisfacer el principio de parsimonia, en el cual se deben expresar las cosas de la manera más simple posible teniendo cuidado de proveer una apropiada concesión a la complejidad causal (p. 10).

El siguiente paso es la construcción de una tabla dicotómica de datos. En la siguiente tabla se establecen con calificación de efectividad (1) a los tres casos en los que se encontró implementación de un número considerable de recomendaciones (considerable se refiere a que al menos una tercera parte de las recomendaciones de la evaluación fueran implementadas). Por otra parte, frente a las variables causales, se agruparon todas las calificaciones intermedias y negativas bajo la etiqueta “0”, y únicamente se dejaron en “1” aquellas variables que

se pueden identificar de manera plena en los procesos de evaluación. En este sentido, depurando variables y mediciones, la tabla de datos dicotómica quedó así:

CASEID	ImpReco	DesCap	Eva	RespPreg	LogInt	ExpMet	MenLimMet	MenDisEv	EvPlazPres
PDAgua	1	1	0	0	1	1	0	0	1
Rep. Colec.	1	1	0	1	1	1	1	1	0
Conta. Aire.	1	0	1	1	0	0	1	0	1
Conexi.intra.	0	0	1	1	1	1	1	0	1
Segu.Alimen.	0	0	1	1	1	0	1	0	1
Defen.Comer.	0	1	1	1	1	1	0	1	1
Retor.Reub.	0	1	1	1	0	1	1	0	0
Aten.huma.	0	1	1	0	1	1	1	0	1

Continúa…

CASEID	IndEv	EvLibAb	InfEv	HaInvUs	AdDisMet	InvUsInt	InvUsRec	ApCaInst
PDAgua	0	0	1	1	1	1	1	1
Rep. Colec.	1	1	1	1	1	1	1	1
Conta. Aire.	0	0	0	1	1	1	1	1
Conexi.intra.	1	1	1	0	0	0	0	1
Segu.Alimen.	0	0	1	0	1	1	0	0
Defen.Comer.	1	1	1	1	1	0	0	1
Retor.Reub.	1	1	1	0	1	1	1	0
Aten.huma.	0	0	0	1	0	1	1	0

La primera variable corresponde a implementación de recomendaciones, es decir, a la efectividad de los procesos evaluativos o el resultado que se busca explicar. Por otra parte, tomando en consideración la recomendación de Rihoux y Ragin acerca de reducir el número de variables a partir de la verificación de que varíen al menos en un tercio de los casos (p. 45), en la aplicación del método comparado con conjuntos nítidos, las únicas variables causales que quedaron (y en su respectivo orden) fueron 16:

- De la OCDE: Desarrollo de capacidad, evaluabilidad, respuesta a preguntas de evaluación, lógica de la intervención, explicación de la metodología, mención de límites metodológicos, mención de discrepancias entre evaluadores, evaluación dentro del plazo y presupuesto previsto, independencia del evaluador, evaluación libre y abierta e informe de la evaluación.
- Del U-FE: Hallar e involucrar a usuarios de la evaluación, adaptación del diseño metodológico, involucramiento de usuarios en la interpretación de

hallazgos, involucramiento de usuarios en la construcción de recomendaciones y apertura al cambio institucional.

Aun así, 16 variables causales continúan siendo un número demasiado elevado. Siguiendo criterios teóricos se pueden seleccionar de este grupo las 8 variables más relevantes para ser integradas al modelo comparativo. Para esto, se seleccionaron las variables que mejor caracterizan las teorizaciones de la OCDE y del U-FE de acuerdo con los criterios por los cuales fueron seleccionados en la presente investigación.

Tal como se mencionó en el primer capítulo, los estándares de la OCDE se seleccionaron a razón de que reflejan de una mejor manera el consenso académico racionalista (ver apartado 1.1.3.2), particularmente en cuanto al énfasis que se realiza en la disponibilidad de información robusta en términos técnicos como factor determinante para el uso de las evaluaciones. Este presupuesto y énfasis se manifiesta de manera evidente dentro de las actuales instituciones que inciden en las situaciones de acción de los procesos de evaluación en Colombia, particularmente a través del rol que ejerce la DSEPP. Por esta razón, las variables de los estándares de calidad de la OCDE seleccionadas son: respuesta a preguntas evaluativas, independencia del evaluador, informe de la evaluación y lógica de la intervención.

Por otra parte, el U-FE fue elegido por desarrollar teóricamente las particulares interacciones que deben existir entre los evaluadores y quienes deben utilizar las evaluaciones. En consonancia con los postulados de Zittoun y Lindblom, el U-FE pone su énfasis en la construcción de acuerdos y ajuste mutuo como base para el desarrollo del apoyo que deben tener las recomendaciones de la evaluación, que dentro de un proceso de (re)formulación pueden categorizarse como propuestas de política pública. En este sentido, las variables más relevantes de ser incluidas dentro del modelo comparativo con: apertura al cambio institucional, hallar e involucrar a usuarios de la evaluación, involucramiento de usuarios en la interpretación de hallazgos e involucramiento de usuarios en la construcción de recomendaciones.

Una vez que se transformaron los datos a una matriz dicotómica, y adicionalmente se redujo el conjunto de variables a partir del proceso la selección aclarado anteriormente, se procedió a utilizar el software fsQCA desarrollado por Charles Ragin[71]. El cual al aplicar la opción de análisis con algoritmo de tabla de verdad generó la siguiente tabla:

[71] El software junto con su manual puede ser encontrados en: http://www.socsci.uci.edu/~cragin/fsQCA/software.shtml

Tabla de verdad – Análisis comparado con conjuntos nítidos

ImpReco	RespPreg	LogInt	IndEv	InfEv	HaInvUs	InvUsInt	InvUsRec	ApCaInst	Number	Raw consist.
1	1	0	0	0	1	1	1	1	1	1
1	0	1	0	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
0	1	1	0	1	0	1	0	0	1	0
0	1	0	1	1	0	1	1	0	1	0
0	0	1	0	0	1	1	1	0	1	0
0	1	1	1	1	0	0	0	1	1	0
0	1	1	1	1	1	0	0	1	1	0

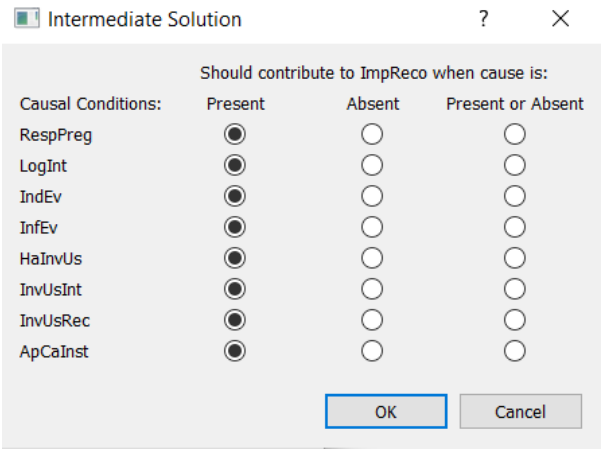
Este ejercicio satisface dos tareas esenciales: valora la distribución de casos entre las diferentes combinaciones lógicas posibles; y valora la consistencia de la evidencia por cada combinación causal bajo el argumento de que los casos con esa combinación constituyen un subconjunto de casos con el resultado esperado (Ragin, 2017, p. 42). La consistencia (columna “Raw consist.”) mide la proporción de casos en una particular combinación causal que presentan el resultado de interés. Cómo se observa en la tabla, y en consonancia con las recomendaciones de Ragin (p. 40), se eliminaron de la tabla de verdad todas aquellas posibles configuraciones causales que no tienen casos que las representen (columna “Number” indica el número de casos con esa combinación causal). No existe ninguna combinación con este particular conjunto de variables causales que tenga más de un caso como evidencia, y únicamente las tres primeras combinaciones tienen consistencia (correspondientes a cada uno de los casos que presentaron resultado positivo).

Luego de este paso, se procede a realizar el análisis estándar dentro del software. Esta opción abre la ventana para escoger los implicantes principales (prime implicants). El concepto de implicación se refiere a que una expresión booleana (como las correspondientes a las combinaciones posibles) implica otra dependiendo si la pertenencia a la segunda es un subconjunto de la pertenencia a la primera. En este sentido, los implicantes primarios son las combinaciones que resultan de usar la regla de minimización booleana (si dos expresiones booleanas difieren únicamente en una condición causal, pero aun así producen el mismo resultado, entonces la condición causal que la distingue puede ser removida para crear una expresión combinada más simple). En la presente investigación, resultaron los siguientes implicantes primarios:

- Involucramiento de usuarios en interpretación de hallazgos y Apertura al cambio institucional (InvUsInt ApCaInst)

- Involucramiento de usuarios en construcción de recomendaciones y Apertura al cambio institucional (InvUsRec ApCaInst)

Que los dos únicos implicantes principales que resultan de esta selección de variables correspondan al U-FE brinda evidencia de la superioridad del U-FE en contra de los estándares de la OCDE, aun así, es necesario completar el análisis. Ambos implicantes principales se mantuvieron en el análisis. Posteriormente, surge la ventana para aclarar los criterios para la solución intermedia (entre la compleja y la parsimoniosa), en donde, de acuerdo con los lineamientos de Ragin, se realizó la siguiente selección:



Esto, a razón de que las teorías aclaran que la relación causal existente, implica que las variables causales deben estar presentes (no ausentes) para que la evaluación sea efectiva. Una vez aclarado lo anterior, el software genera como salida tres tipos de soluciones distintas: la compleja, la parsimoniosa y la intermedia:

Solución compleja:

```
                                                          raw        unique
                                                         coverage   coverage   consistency
                                                         ---------- ---------- ----------
RespPreg*~LogInt*~IndEv*~InfEv*HaInvUs*InvUsInt*InvUsRec*ApCaInst     0.333333   0.333333   1
~RespPreg*LogInt*~IndEv*InfEv*HaInvUs*InvUsInt*InvUsRec*ApCaInst      0.333333   0.333333   1
RespPreg*LogInt*IndEv*InfEv*HaInvUs*InvUsInt*InvUsRec*ApCaInst        0.333333   0.333333   1
solution coverage: 1
solution consistency: 1
```

Solución parsimoniosa:

	raw coverage	unique coverage	consistency
InvUsInt*ApCaInst	1	0	1
InvUsRec*ApCaInst	1	0	1

solution coverage: 1
solution consistency: 1

Solución intermedia:

	raw coverage	unique coverage	consistency
RespPreg*HaInvUs*InvUsInt*InvUsRec*ApCaInst	0.666667	0.333333	1
LogInt*InfEv*HaInvUs*InvUsInt*InvUsRec*ApCaInst	0.666667	0.333333	1

solution coverage: 1
solution consistency: 1

Para poder interpretar estos resultados se debe explicar la diferencia que existe entre estos tres tipos de soluciones. En primer lugar, teniendo en cuenta la regla de minimización booleana se debe decir que la solución intermedia es un subconjunto de la solución parsimoniosa y un superconjunto de la solución compleja. Como se puede evidenciar, las variables que se incluyen en la solución intermedia son menos numerosas que las que se incluyen en la solución compleja, pero mayor en cantidad de las que se incluyen en la solución parsimoniosa.

Esto es así, porque para la solución compleja el análisis excluye el uso de contrafácticos[72], mientras que la solución parsimoniosa incluye contrafácticos tanto de las configuraciones causales con consistencia, como de las que no tuvieron consistencia (p. 44-45). Por su parte, la solución intermedia resulta cuando se incorporan a los resultados diferentes subconjuntos de los residuos utilizados para producir la solución parsimoniosa, es decir, son soluciones que utilizan al menos algunas de las condiciones causales especificadas en la solución compleja y que tienen todas las condiciones causales especificadas en la solución parsimoniosa (p. 45).

Adicionalmente, la solución compleja ofrece la solución más detallada en el sentido en que asume que todas las configuraciones sin casos (number=0, también

[72] Los contrafacticos en el QCA son sustitutos de los grupos de control en los experimentos. Puntualmente se refieren a casos hipotéticos que se generan a partir de invertir los valores (binarios) del resultado y de las variables causales (Ragín, 2017, p. 43). Por ejemplo, si AB produce Y, entonces la negación de A y la Negación de B producen la Negación de Y. Así como la tabla de verdad apoya en la identificación de patrones transversales a los casos, los contrafactuales también generan sus propios patrones.

llamadas “remainders”) hubiesen producido la ausencia del resultado de interés (todos los “remainders” se configuran como falsos), lo que impide la simplificación de las configuraciones. Mientras que la solución parsimoniosa ofrece la solución más simple, a razón de que utiliza a cualquier configuración causal sin casos que ayude a generar una solución lógicamente más simple (p. 56). Por su parte, la solución intermedia supone que sólo algunas configuraciones sin casos hubiesen producido el resultado de interés dependiendo de que utilicen contrafácticos “fáciles”[73].

En las soluciones “raw coverage” indica la proporción de casos con resultado positivo que pueden ser explicados a partir de esa configuración causal. La columna de “unique coverage” indica la proporción de casos con resultado positivo que son explicados exclusivamente por las configuraciones. La columna consistencia, como se mencionó, indica la proporción de casos positivos que presentan esa configuración causal. “Solution coverage” indica la proporción de casos con resultados positivos que son explicados con el modelo final, y “solution consistency” indica la proporción de casos incluidos en la solución que presentan el resultado de interés. Las tres soluciones presentan una proporción de 1 en estos dos indicadores.

Estos resultados nos producen dos posibles interpretaciones. En primer lugar, el hecho de que las dos configuraciones causales que aparecen en la solución parsimoniosa incluyen únicamente variables del U-FE (apertura al cambio institucional, involucramiento de usuarios en la interpretación de resultados, e involucramiento de usuarios en la construcción de recomendaciones), indica que el U-FE es superior a los estándares de calidad de la OCDE para explicar el uso de la evaluación en Colombia. Sin embargo, este resultado falsea la hipótesis central de este estudio por lo menos de manera parcial, puesto que la efectividad de la evaluación únicamente dependería de la coordinación entre el proceso de evaluación y el proceso de política pública, y no tendría sustento la primera variable independiente de la hipótesis de investigación.

En segundo lugar, tanto la solución compleja como la intermedia a parte de incluir a todas las variables del U-FE dentro de las configuraciones causales con consistencia y cobertura relevante, también incluyen las variables de estándares de calidad: Respuesta a preguntas evaluativas, Lógica de la intervención e Informe de

[73] La diferencia entre un contrafáctico “fácil” o” difícil” radica en que los primeros se apoyan en el conocimiento empírico existente, mientras que los segundos no (Ragin, 2017, p. 45). También puede decirse que la designación de "fácil" versus "Difícil" se basa en la información proporcionada por el usuario con respecto a la conexión entre cada condición causal y el resultado (p. 56).

la evaluación. La variable de independencia del evaluador es descartada en la solución intermedia, y no solo es considerada en una de las configuraciones causales posibles de la solución compleja (mientras que las otras tres relacionadas con calidad son incluidas en dos configuraciones). Adicionalmente, en la solución intermedia se observa que el resultado se puede dar por la multiplicación entre todas las variables del U-FE por la respuesta a preguntas evaluativas, o por la multiplicación entre lógica de la intervención e informe de la evaluación. Estas soluciones privilegian las variables procedentes del U-FE por encima de las de la OCDE, sin embargo, no descartan a este último grupo, lo que confirmaría la hipótesis central.

12. Conclusión

El estudio de caso a profundidad de la política pública de evaluación de políticas que fue realizado en el capítulo 2, generó una serie de conclusiones descriptivas preliminares que vale la pena resaltar en este punto. Esto con el propósito de valorar si la información que se obtuvo en los capítulos correspondientes a los casos a comparar, y en el análisis comparado, valida o rechaza esas conclusiones preliminares. A continuación, se presenta esta información:

- La institucionalidad de la política pública de evaluación de políticas le brinda al DNP múltiples competencias que en la literatura académica son asignadas a los evaluadores. Al mismo tiempo, al ser el DNP quien contrata y supervisa la evaluación, existen riesgo de que la evaluación esté demasiado orientada a satisfacer las preferencias del DNP en detrimento de las necesidades de los usuarios de la evaluación en los programas evaluados. Esta situación se reflejó en los casos comparados (particularmente en las evaluaciones de las medidas de defensa comercial y de conexiones intradomiciliarias).
- Se manifestó que no hay requerimientos institucionales puntuales acerca de la utilización prevista de las evaluaciones. Tampoco se encontró evidencia acerca de discusiones alrededor del uso de la evaluación. Esta conclusión se debe matizarse, puesto que en el análisis a profundidad de los casos se encontró que los requerimientos del plan de transferencia e implementación de recomendaciones implica orientar la evaluación a su uso desde la etapa de contratación (a pesar de que estos planes no lleguen a implementarse). Sin embargo, lo anterior no basta para poder decir que hay una plena claridad acerca del uso que se le va a dar a las evaluaciones por parte de los actores participantes.
- Las instituciones orientadas a satisfacer los estándares de control de calidad y de inclusión de observaciones de implicados pueden fácilmente entrar en conflicto con el estándar de independencia del evaluador, reflejándose esto en manifestación de gerencialismo. Esta conclusión preliminar se confirma. A pesar de que siempre se cumplen los estándares de control de calidad y de inclusión de observaciones, únicamente en la mitad de los procesos se cumplió el estándar de independencia del evaluador, el cual se vio afectado

por requerimientos de cambios substanciales en los juicios y conclusiones de la evaluación.

- No existe claridad sobre quiénes son los usuarios primarios de las evaluaciones, lo que genera un déficit de capacidad institucional para implementar las recomendaciones. Esta conclusión del análisis descriptivo se confirma, en efecto, se encontró que muchas veces se diseñan recomendaciones cuya implementación depende de actores que no están involucrados en el proceso de evaluación.
- Se planteó que el elevado grado de control de la DSEPP incide de manera negativa en la fluidez de los ajustes mutuos que se deben dar entre el evaluador y los usuarios de la evaluación. A pesar de las barreras en el lenguaje, y el elevado grado de tecnicismo, esta conclusión se debe descartar. El establecimiento del comité de diseño de la evaluación y del comité de seguimiento fueron reconocidos, en la mayoría de los casos, como escenarios efectivos para la comunicación de las preocupaciones y demandas de los usuarios.
- Las evaluaciones tienen una naturaleza rígida sin capacidad de adaptarse a las demandas cambiantes de los usuarios de las evaluaciones, lo que genera pérdida de relevancia de los resultados de la evaluación. Las mediciones de la efectividad de las evaluaciones y de la apertura al cambio institucional confirman en varios casos esta afirmación.
- No se mide el uso de la evaluación, no se le realiza seguimiento, y a pesar de que la DSEPP toma el rol de evaluador frente a muchas de las responsabilidades que este debería tener, no pasa igual frente a la apropiación y defensa activa de los resultados de las evaluaciones. Esta afirmación se confirma con los datos de los casos, únicamente en un PTIR se consideraron indicadores de seguimiento para implementar las recomendaciones. En general, se observa que la relación entre la DSEPP y los usuarios de las evaluaciones finaliza una vez se difunden los resultados de las evaluaciones.

En cuanto al análisis comparado, en primer lugar, debe recordarse la relevancia teoría que fue aclarada en la presente investigación. Como se mencionó en el primer capítulo, la presente tesis busca poner a prueba o evaluar premisas no cuestionadas en gran parte de la bibliografía (King, Verba y Keohane, 2000, p. 27). También se aclaró que los estándares de calidad de la OCDE fueron seleccionados a razón de que, como herramienta de metaevaluación, reflejan institucionalmente el consenso académico racionalista.

La corriente racionalista que fue presentada como uno de los enfoques que se encuentran en la literatura en cuanto a las explicaciones del uso de las evaluaciones, se enfoca en hipótesis cómo: la realización de evaluaciones crea las condiciones para la eficiencia a partir de reducir asimetrías en la información y los costos de transacción y mejorando el flujo de información (Johanson y Siverbo,

2009); la medición del desempeño a través de indicadores viene acompañada invariablemente de actividades de uso de la información para la toma de decisión, y el uso de la información para la toma de decisiones puede ser facilitado principalmente a partir de avances institucionales y técnicos en las mediciones (Taylor, 2011, p. 1).

Este enfoque se caracteriza por tener un mismo supuesto del neoinstitucionalismo económico de North, particularmente en el sentido en que los actores sociales buscan maximizar sus objetivos dentro de un marco institucional dado, para lo cual, invariablemente van a utilizar la información que tengan a su disposición. Así mismo, los estándares de calidad de la OCDE plasman su énfasis en las condiciones necesarias para que la información producida por la evaluación esté disponible, sea rigurosa y acertada.

A pesar de que los estándares de calidad no dejan por fuera temas como: participantes relevantes, coordinación o inclusión de observaciones de implicados (que pueden ser considerados como similares a los del U-FE), su conceptualización es bastante precaria. La participación de los demás actores se restringe a la etapa de diseño de la evaluación; la coordinación, a tener en cuenta sistemas, planes, actividades y políticas nacionales y locales en la materia; y la inclusión de observaciones se restringe al informe final de la evaluación.

Esto significa, que los estándares de la OCDE tienen el mismo supuesto del racionalismo y el neoinstitucionalismo económico: los actores solo necesitan tener la información disponible para usarla, puesto que van a buscar maximizar sus objetivos con esta. Y esto se ve plasmado claramente en la institucionalidad de la política pública de evaluación de políticas: la relación termina en el momento en que se socializan los resultados de la evaluación.

Por otro lado, el U-FE fue elegido porque ayuda a poner a prueba esa premisa no lo suficientemente discutida y plasmada en la institucionalidad colombiana. El U-FE está en armonía con los postulados de Zittoun y Lindblom, en el sentido en que, si se quiere que una evaluación sea usada, debe partirse de considerar sus recomendaciones como (re)formulaciones de política pública, las cuales, para desarrollar algún apoyo (e interiorización), deben apoyarse en la construcción de acuerdos y ajustes mutuos, traducidos como coordinación entre el equipo evaluador y los usuarios de la evaluación.

Las inferencias descriptivas que permitió obtener el método de diferencia fueron: no existe ninguna variable causal tanto del modelo OCDE como del U-FE que varíe igual a la variable de efectividad, por lo que ninguna variable puede ser considerada como independientemente suficiente para explicar el cambio en la variable independiente; el caso en el que más hubo estándares de calidad insatisfechos fue el caso con más recomendaciones implementadas; y la combinación del hallazgo e involucramiento de los usuarios de la evaluación, la reactividad del DNP y del equipo

evaluador ante las preocupaciones e inquietudes de los usuarios en las etapas de definición del uso, las preguntas, metodología, adaptación de metodología, interpretación de hallazgos, construcción de recomendaciones, siempre se presenta al tiempo que existe la efectividad de las evaluaciones.

Esta última inferencia fue desafiada a partir del caso de la evaluación de la Política de Atención Humanitaria, la cual presentaba todos esos factores causales del U-FE, pero no presentaba efectividad. Esto hizo surgir preguntas acerca de las diferencias que tenían los casos efectivos con el caso de la evaluación de la Política de Atención Humanitaria, por lo que, al profundizar en el análisis de los datos recolectados, y al identificar variables el U-FE en las cuales esto pudiera ser categorizado, se encontró que se debían realizar mediciones complementarias en todos los casos a partir de la nueva variable: apertura al cambio institucional por parte de los usuarios de la evaluación.

Las nuevas mediciones permitieron obtener que la combinación: hallar e involucrar usuarios de la evaluación, escuchar preocupaciones y demandas de usuarios, trabajo conjunto sobre usos, discusión de debates metodológicos, adaptación del diseño metodológico a necesidades cambiantes de usuarios, involucramiento de usuarios en interpretación de hallazgos, involucramiento de usuarios en la construcción de recomendaciones y la apertura al cambio institucional por parte de usuarios, es una combinación que varía de igual manera que la variable dependiente.

Sin embargo, a pesar de que la lógica del método de diferencia permitió obtener la anterior inferencia preliminar, el método de diferencia no tiene la capacidad de generar inferencias conclusivas alrededor de combinaciones causales que son suficientes para explicar el cambio de una variable dependiente. Si bien esto no permitió obtener las conclusiones necesarias con relación a la hipótesis y el objetivo de investigación o a la relevancia teórica de la misma, si permitió que surgiera la necesidad de realizar un análisis comparado con mayor profundidad.

Para avanzar hacia conclusiones acerca de un conjunto de variables que puedan explicar la variación de la efectividad de las evaluaciones de políticas públicas, fue necesario utilizar el método comparado con conjuntos nítidos desarrollado por Ragin. Para esto, se debió partir de una adecuada selección de variables, para lo cual, se descartaron las variables causales que no tienen suficiente variabilidad y que, por ende, pueden ser consideradas como constantes. Posteriormente, del conjunto de 16 variables que quedaron se seleccionaron las 8 más representativas teniendo consideración de los criterios teóricos por los cuales fueron en un principio seleccionados los enfoques de la OCDE y el U-FE.

Luego se procedió a construir una tabla dicotómica de datos en correspondencia del algebra booleana que es utilizada en este método. Para esto, todas las calificaciones intermedias entre la plena existencia de la variable y la no existencia

de esta, fueron transformadas en una calificación negativa, de tal manera que las variables con medición positiva en los casos únicamente correspondieran a manifestaciones plenas de los fenómenos que pretenden medir.

Posteriormente, se construyó la tabla de verdad, que luego de su depuración de configuraciones causales posibles que no tuvieran ningún caso que las representara, generó 8 configuraciones causales a considerar en el análisis. Una vez aclarado el modo en el que las variables contribuyen teóricamente a la existencia de la variable dependiente (con su presencia), y que se seleccionaran los dos únicos implicantes principales resultantes, se obtuvieron las tres soluciones brindadas por el método: solución compleja, intermedia y parsimoniosa. Este análisis generó las siguientes inferencias:

- Los dos únicos implicantes principales que resultaron del ejercicio corresponden a: Involucrar a usuarios en interpretación de hallazgos*Apertura de usuarios al cambio institucional; e Involucrar a usuarios en la construcción de recomendaciones*Apertura de usuarios al cambio institucional. Lo que cual brinda evidencia a favor de la superioridad del U-FE ante los estándares de la OCDE como mejor explicación de la variabilidad de la efectividad de los procesos de evaluación. Esto contribuye a la relevancia teórica que fue seleccionada para la presente investigación.
- Las dos configuraciones causales que aparecen en la solución parsimoniosa corresponden a los dos implicantes principales mencionados en el punto anterior. La solución parsimoniosa es la solución más simple y general que produce la aplicación del algoritmo de tabla de verdad con conjuntos nítidos. Esto también le brinda fuerza a la inferencia de que el U-FE es superior a los estándares de la OCDE para explicar la efectividad de los procesos de evaluación.
- Tanto la solución compleja como la solución intermedia incluyen configuraciones causales en las que están todas las variables seleccionadas del U-FE. En la solución intermedia se descarta dentro de las configuraciones resultantes a la independencia del evaluador como variable causal. Adicionalmente, se propone que para que exista efectividad de la evaluación deben darse las cuatro variables del U-FE con dos opciones: Respuesta a preguntas evaluativas; o Lógica de la intervención*Informe de la evaluación. Esto también brinda fuerza a la inferencia de la superior del U-FE, sin embargo, hace la salvedad de que se deben incluir factores relacionados con los estándares de la OCDE, a pesar de que estos pueden ser intercambiables entre sí.

Estas inferencias permiten establecer que la hipótesis central de investigación: la satisfacción de los estándares profesionales de calidad en los procesos de evaluación y la coordinación entre los procesos de evaluación y de las políticas públicas evaluadas generan la efectividad de la política pública de evaluación de políticas en Colombia, no es verdadera. La satisfacción de los estándares

profesionales de calidad no es una variable esencialmente necesaria o suficiente para que exista la efectividad de la política de evaluación medida a partir del uso de la información.

Sin duda alguna, los estándares de calidad sirven para tener información con capacidad de ser utilizada, pero el uso de la evaluación no depende de la existencia o no de este tipo de información. Se puede afirmar que el uso de la evaluación, traducido en la implementación de recomendaciones como manifestación de un mejoramiento en la toma de decisiones, se relaciona mucho más con la forma en que se coordinan los actores participantes en la evaluación, en el modo en que los usuarios de la evaluación son involucrados en formas que consideran relevantes, y particularmente en la voluntad que tengan de modificar la institucionalidad de las políticas evaluadas teniendo como insumo los hallazgos de la evaluación.

En este sentido, y teniendo en consideración el objetivo central de esta investigación, se puede afirmar que la política pública de evaluación de políticas en Colombia genera información relevante para poder generar evidencia acerca de las teorizaciones de la OCDE y de Patton. Particularmente brinda evidencia en contra de la importancia que tienen los estándares de calidad para una efectividad de la evaluación, mientras que genera evidencia a favor de la importancia que tiene una adecuada coordinación entre el proceso de evaluación con el de política pública evaluada.

Finalmente, se brindan algunas recomendaciones para mejorar la actual política de evaluación de políticas en Colombia:

- Medir su efectividad a partir del uso de las evaluaciones que producen por parte de las entidades sectoriales que evalúan.
- Evitar fundamentar los desarrollos institucionales (ya sean jurídicos o técnicos) en el supuesto de que a mayor grado rigurosidad técnica y científica de la información, más probable será el uso de la evaluación.
- Realizar desarrollos institucionales orientados a la utilización de la evaluación, particularmente a la implementación de los PTIR. Para esto, es necesario que la DSEPP realice un seguimiento efectivo a la implementación de las recomendaciones.
- Enfocar posteriores desarrollos institucionales en el fortalecimiento de la coordinación del proceso de evaluación con el de la política pública evaluada. Particularmente, es necesario identificar e involucrar a todos los usuarios de la evaluación (así esto implique modificar los participantes en los Comités); caracterizar la apertura al cambio institucional de estos usuarios, priorizando recomendaciones hacia quienes tengan esta apertura; adaptar la evaluación a las necesidades cambiantes de los usuarios; requerir a los contratistas la elaboración de informes preliminares sobre hallazgos que reduzcan las sorpresas y aumenten la oportunidad de la información; y garantizar la

participación de los usuarios en la interpretación de hallazgos y en la construcción de recomendaciones (sin afectar la independencia del evaluador).

- Es necesario trascender el supuesto de que la implementación de las recomendaciones es discrecionalidad de las entidades sectoriales evaluadas. La DSEPP debe apropiarse de los resultados de las evaluaciones que produce, comprendiendo que las recomendaciones son propuestas de política pública, y como tal, es necesario que estas sean difundidas y defendidas en los escenarios que requieran.

Bibliografía

- Álvarez, Ángel Eduardo (1992) Análisis de políticas públicas. CLAD. Serie Temas de Coyuntura en Gestión Pública. Caracas.
- Ammons, David y Rivenbark, William (2008). Factors Influencing the Use of Performance Data to Improve Municipal Services: Evidence from the North Carolina Benchmarking Project. Public Administration Review, 68: 304-318.
- Appleton-Dyer, Sarah; Clinton, Janet; Carswell, Peter y McNeill, Rob (2012). Understanding Evaluation Influence Within Public Sector Partnerships: A Conceptual Model. American Journal of Evaluation, 33 (4).
- Askim, Jostein (2007). How do politicians use performance information? An analysis of the Norwegian local government experience. International Review of Administrative Sciences, 73.
- Balthasar, Andreas (2009). Institutional Design and Utilization of Evaluation. A Contribution to a Theory of Evaluation Influence Based on Swiss Experience. Evaluation Review, 33 (3).
- Banco Mundial (2014). Invitación a presentar expresiones de interés (servicios de consultoría-selección de firmas) BM-008-2014. Proyecto Fortalecimiento de la Información Pública Seguimiento y Evaluación para la Gestión por Resultados en Colombia. Evaluación de Atención Humanitaria. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do
- Berman, Evan y Wang, Xiaohu (2002). Performance Measurement in U.S. Counties: Capacity for Reform. Public Administration Review 60 (5).
- Borras, Susana y Hojlung, Steven (2015). Evaluation and policy learning: The learners' perspective. En: European Journal of Political Research 54: 99-120.
- Bourdeaux, Carolyn y Chikoto, Grace (2008). Legislative influences on performance management reform. Public Administration Review, 68
- Boussaguet, Laurie, Jacquot, Sophie, Ravinet, Pauline (2010). Dictionnaire des politiques publiques. Presses de la FNSP. Paris.
- Campbell, John, Pedersen, Ove (2001). The Rise of Neo-liberalism and Institutional Analysis. Princeton University Press. Princeton.
- Centro Nacional de Consultoría (2010a). Línea de base del acceso, uso y la calidad de la información que produce SINERGIA. DNP. Bogotá. Colombia.
- _____ (2010b). Ficha técnica: Línea de base del acceso, uso y la calidad de la información que produce SINERGIA. DNP.

Bogotá. Colombia. Recuperada en: https://sinergiapp.dnp.gov.co/#Evaluaciones/EvalFin/145

- _____ (2015a) Ficha Técnica: Evaluación de impacto y de diseño del Programa Conexiones Intradomiciliarias de Ministerio de Vivienda, Ciudad y Territorio. DNP. Bogotá. Colombia. Recuperada en: https://sinergiapp.dnp. gov.co/#Evaluaciones/EvalFin/145
- _____ (2015b) Evaluación de impacto y de diseño del programa conexiones intradomiciliarias del Ministerio de Vivienda, Ciudad y Territorio. Informe final (producto 4). DNP. Bogotá. Colombia.
- _____ (2015c) Plan de Transferencia e Implementación de Recomendaciones. Evaluación de impacto y de diseño del Programa de Conexiones Intradomiciliarias. DNP. Bogotá. Colombia.
- _____ (2016a) Evaluación institucional y de resultados del Programa de Reparación colectiva (Ficha Técnica). DNP. Bogotá Colombia.
- _____ (2016b) Producto 4: Informe de resultados de la evaluación. Realizar una evaluación institucional y de resultados del Programa de Reparación Colectiva, que permita identificar las acciones a mejorar sobre la articulación y la capacidad institucional de las entidades a cargo de su formulación e implementación, a nivel nacional y territorial, para el cumplimiento de sus objetivos de acuerdo a lo establecido en el Decreto 4800 de 2011. DNP. Bogotá. Colombia.
- _____ (2016c). Producto 5: Plan de Transferencia e Implementación de Recomendaciones – PTIR. Realizar una evaluación institucional y de resultados del Programa de Reparación Colectiva. DNP. Bogotá. Colombia.
- _____ (2016d). Resumen ejecutivo de la evaluación (evaluación institucional y de resultados del Programa de Reparación Colectiva. DNP. Bogotá. Colombia.
- De Lancer J.; Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: an empirical study of factors affecting adoption and implementation. Public Admin Review. 61. 693-708.
- Denzau, Arthur y Douglass North (2000). Shared Mental Models: Ideologies and Institutions. En: Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin. Elements of Reason—Cognition, Choice, and the Bounds of Rationality, ed. Cambridge University Press. Cambridge.
- Departamento Administrativo para la Prosperidad Social (2018). –-Borrador- Decreto Número de 2017 “Por medio del cual se modifican y adicionan algunos artículos al Decreto 1084 de 2015 Único Reglamentario del sector de Inclusión Social y Reconciliación”. Documento de trabajo.
- DNP (2009). Guía metodológica para la formulación de indicadores. Editorial Scripto Gómez y Rosales Asociados Ltda. Bogotá. Colombia.
- ____ (2012). Aviso de convocatoria pública. Solicitud de expresiones de interés para la conformación de la lista corta del concurso de méritos N° CM-026-2012. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do

- ____ (2014a). Manual de Contratación. DNP. Bogotá. Colombia.
- ____ (2014b). Contrato de consultoría celebrado entre el Departamento Nacional de Planeación y Centro Nacional de Consultoría N° 512 de 2014. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do
- ____ (2014c). Contrato de consultoría celebrado entre el Departamento Nacional de Planeación y la Unión Temporal Econometría S.A. -Sistemas Especializados de Información S.A. No. DNP-OR-056-2014. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do
- ____ (2014d). Aviso de Convocatoria Proceso de Selección por Concurso de Méritos No. CM-023-2014. DNP. Bogotá. Colombia. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do
- ____ (2014e). Contrato de consultoría celebrado entre el Departamento Nacional de Planeación y G/Exponencial consultores SAS. No. DNP-509-2014. Recuperado de https://www.icbf.gov.co/sites/default/files/pnsan.pdf
- ____ (2015a). Plan Nacional de Desarrollo: Todos por un Nuevo país 2014-2018. DNP. Bogotá. Colombia.
- ____ (2015b). Contrato de Consultoría celebrado entre el Departamento Nacional de Planeación y el Centro Nacional de Consultoría S.A. No. DNP-827-2015. Recuperado de: https://www.contratos.gov. co/consultas/ inicioConsulta.do
- ____ (2016a). Aviso de convocatoria proceso de selección concurso de méritos No. CM-007-16. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do
- ____ (2016b). Contrato de Consultoría celebrado entre el Departamento Nacional de Planeación y UT Econometría S.A.-SEI [CM-007-2016] No. DNP-572-2016. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do
- Dryzek, John S. (2000). Deliberative Democracy and Bey. Oxford University Press. Oxford. Reino Unido.
- DSEPP-DNP (2012a). Guía para la evaluación de políticas públicas. DNP. Bogotá. Colombia.
- _____ (2012b). Guía Metodológica para el Seguimiento al Plan Nacional de Desarrollo y la Evaluación de Políticas Estratégicas. DNP. Bogotá. Colombia.
- _____ (2012c). Estudios previos contratación por concurso de méritos con precalificación (proceso: No. DNP-OR-056-2014). DNP. Colombia. Recuperado de: https://www.contratogov.co/ consultas/inicioConsulta.do
- _____ (2014a). Guía metodológica para el seguimiento y la evaluación a políticas públicas. DNP. Colombia.
- _____ (2014b). Análisis del sector de consultoría en evaluación de políticas públicas. DNP. Bogotá. Colombia.
- _____ (2014c). Proceso de selección por concurso de méritos abierto No. CM-021-14. Pliego de Condiciones Definitivo. objeto: Realizar una evaluación de impacto del Programa Conexiones Intradomiciliarias, que permita medir el efecto de la intervención en las variables de resultado relacionadas con el desarrollo humano y el mejoramiento de la calidad de vida de la población

beneficiada. Recuperado de: https://www.contratogov.co/consultas/inicioConsulta.do

- _____ (2014d). Proceso de selección por concurso de méritos abierto No. CM-020-14. Pliego de Condiciones. Objeto: realizar una evaluación institucional y de resultados del programa “Planes Departamentales para el manejo Empresarial de los Servicios de Agua y Saneamiento”. Recuperado de: https://www.contratos.gov.co/consultas/inicioConsulta.do
- _____ (2014e). Concepto técnico Producto N° 3 Informe de resultados y segunda entrega de la documentación de las bases de datos de la evaluación de la Política de Atención Humanitaria a Población Desplazada. DNP. Bogotá. Colombia.
- _____ (2014f). Concepto técnico Producto N° 4 Plan de Transferencia e Implementación de Recomendaciones y tercera entrega de la documentación de las bases de datos de la evaluación de la Política de Atención Humanitaria a la Población Desplazada. DNP. Bogotá. Colombia.
- _____ (2014g). Proceso de Selección por Concurso de Méritos Abierto No. CM-023-14. Pliego de Condiciones Definitivo. Objeto: Realizar una evaluación institucional y de resultados de la Política Nacional de Seguridad Alimentaria y Nutricional –PSAN-. Recuperado de: https://www.contratos.gov.co/consultas/inicioConsulta.do
- _____ (2015a). Concepto técnico Producto N° 4 – Informe Final Evaluación de Impacto del Programa Conexiones Intradomiciliarias. DNP. Bogotá. Colombia.
- _____ (2015b). Concepto técnico Producto N° 5 – Informe Plan de Transferencia e Implementación de Recomendaciones Evaluación de Impacto del Programa Conexiones Intradomiciliarias. DNP. Bogotá. Colombia.
- _____ (2015c). Concepto técnico Producto N° 5 – Informe Plan de Transferencia e Implementación de Recomendaciones Evaluación Institucional y de Resultados del Programa Planes Departamentales de Agua. DNP. Bogotá. Colombia.
- _____ (2015d). Concepto técnico Producto N° 4 – Informe de Resultados Evaluación Institucional y de Resultados del Programa Planes Departamentales de Agua. DNP. Bogotá. Colombia.
- _____ (2015e). Proceso de Selección por Concurso de Méritos Abierto No. CM-013-15. Pliego de Condiciones Definitivo. OBJETO: Realizar una evaluación institucional y de resultados del Programa de Reparación Colectiva, que permita identificar las acciones a mejorar sobre la articulación y la capacidad institucional de las entidades a cargo de su formulación e implementación, a nivel nacional y territorial, para el cumplimiento de sus objetivos de acuerdo a lo establecido en el Decreto 4800 de 2011. Recuperado de: https://www.contratos.gov.co/consultas/inicioConsulta.do
- _____ (2015f). Concepto técnico producto No 4 Informe de resultados y tercera entrega de la documentación de las bases de datos de la evaluación de la Política de Seguridad Alimentaria y Nutricional. DNP. Bogotá. Colombia.

- _____ (2015g). Concepto técnico producto No. 5 Plan de Transferencia e Implementación de Recomendaciones – PTIR y cuarta entrega de la documentación de las bases de datos de la evaluación de la Política de Seguridad Alimentaria y Nutricional. DNP. Bogotá. Colombia.
- _____ (2015h). Proceso de selección por Concurso de Méritos Abierto No. CM-017-15. Pliego de Condiciones Definitivo. Objeto: Realizar una evaluación de operaciones que permita identificar las fortalezas y debilidades de las actividades que se ejecutan en el proceso de investigación e implementación de una medida de defensa comercial (salvaguardia y derecho antidumping), con el fin de generar recomendaciones conducentes a mejorar la calidad y pertinencia del mismo. Recuperado de: https://www.contratos.gov.co/consultas/inicioConsulta.do
- _____ (2016a). Proceso de selección por concurso de méritos abierto No. CM-007-16. Pliego de Condiciones Definitivo. Objeto: Realizar una evaluación de operaciones y de resultados de la política de retornos y reubicación que permita identificar las fortalezas y debilidades de las acciones de atención a las víctimas de desplazamiento forzado; así como la sostenibilidad de los procesos de retorno y reubicación. Recuperado de: https://www.contratos.gov.co/consultas/inicioConsulta.do
- _____ (2016b). Concepto técnico producto No. 4 Informe de Resultados de la Evaluación (Programa de Reparación Colectiva. DNP. Bogotá. Colombia.
- _____ (2016c). Concepto técnico producto No. 5 Plan de Transferencia e Implementación de Recomendaciones de la Evaluación (Programa de Reparación Colectiva. DNP. Bogotá. Colombia.
- _____ (2016d). Concepto técnico Producto N° 3 Informe Final Evaluación de operaciones a las medidas de defensa comercial (salvaguardias y derecho antidumping). DNP. Bogotá. Colombia.
- _____ (2016e). Concepto técnico producto No. 4 Plan de Transferencia e Implementación de Recomendaciones de la evaluación de operaciones de las medidas de defensa comercial (salvaguardias y derecho antidumping). DNP. Bogotá. Colombia.
- _____ (2016f). Proceso de selección por Concurso de Méritos Abierto No. CM-018-16. Pliego de Condiciones Definitivo. Realizar una evaluación de resultados de la Política de Prevención y Control de la Contaminación del Aire. Recuperado de: https://www.contratos.gov.co/consultas/inicioConsulta.do
- _____ (2016g). Concepto Técnico Producto No. 1 Informe Metodológico evaluación de resultados de la Política de Prevención y Control de la Contaminación del Aire (PPCCA). DNP. Bogotá. Colombia.
- _____ (2016h). Contrato de consultoría celebrado entre el Departamento Nacional de Planeación y G/Exponencial Consultores SAS No. DNP-494-2016. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do
- _____ (2016i). Contrato de consultoría celebrado entre el Departamento Nacional de Planeación y Economía Urbana Limitada No. DNP-605-2016. Recuperado de: https://www.contratos.gov.co/consultas/ inicioConsulta.do

- _____ (2017a). Concepto técnico producto No. 4 Informe de Resultados de la Evaluación y tercera entrega de la documentación de las bases de datos. Política de Retornos y Reubicaciones. DNP. Bogotá. Colombia.
- _____ (2017b). Concepto técnico producto No. 5 Plan de Transferencia e Implementación de Recomendaciones- PTIR y tercera entrega de la documentación de las bases de datos. Política de Retornos y Reubicaciones. DNP. Bogotá. Colombia.
- _____ (2017c). Concepto técnico producto No. 1 Informe metodológico de la evaluación de la política de retornos y reubicaciones. DNP. Bogotá. Colombia.
- _____ (2017d). Concepto Técnico Producto No. 3 Plan de Transferencia e Implementación de Recomendaciones (PTIR) de la evaluación de resultados de la Política de Prevención y Control de la Contaminación del Aire (PPCCA). DNP. Bogotá. Colombia.
- _____ (2017e). Concepto Técnico Producto No. 4 Informe de Resultados de la evaluación de la Política de Prevención y Control de la Contaminación del Aire (PPCCA). DNP. Bogotá. Colombia.
- _____ (sf). Evaluación de Operaciones. DNP. Bogotá. Colombia. Recuperado de: https://sinergia.dnp.gov.co/Paginas/ Internas/Evaluaciones/Metodologías.aspx
- _____ (sf) .Evaluación Institucional. DNP. Bogotá. Colombia. Recuperado de: https://sinergia.dnp.gov.co/Paginas/ Internas/Evaluaciones/Metodologías.aspx
- _____ (sf). Evaluación de Resultados. DNP. Bogotá. Colombia. Recuperado de: https://sinergia.dnp.gov.co/Paginas/ Internas/Evaluaciones/Metodologías.aspx
- _____ (sf). Evaluación de Impacto. DNP. Bogotá. Colombia. Recuperado de: https://sinergia.dnp.gov.co/Paginas/ Internas/Evaluaciones/Metodologías.aspx
- Econometría y Sistemas Especializados de Información (2014a). Evaluación de Operaciones de la Política de Atención Humanitaria a Población en situación de Desplazamiento. Producto 3: Informe de resultados. DNP. Bogotá. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2014b). Evaluación de Operaciones de la Política de Atención Humanitaria a Población en situación de Desplazamiento. Ficha técnica. DNP. Bogotá. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2014c). Evaluación de Operaciones de la Política de Atención Humanitaria a Población en Situación de Desplazamiento. Producto 4: Plan de Transferencia e Implementación de Recomendaciones y tercera entrega de la documentación de las bases de datos de la evaluación. DNP. Bogotá. Colombia.
- _____ (2017a). Evaluación de operaciones y de resultados de la Política de Retornos y Reubicaciones. Producto 4: Informe de Resultados de la Evaluación. DNP. Bogotá. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145

- _____ (2017b). Evaluación de operaciones y de resultados de la Política de Retornos y Reubicaciones. Resumen Ejecutivo. DNP. Bogotá. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2017c). Evaluación de operaciones y de resultados de los Procesos de Retornos y Reubicaciones. Producto 5: Plan de Transferencia e Implementación de las Recomendaciones. DNP. Bogotá. Colombia.
- Economía Urbana (2017a). CTTO DNP-605 de 2016 Realizar una evaluación de resultados de la Política de Prevención y Control de la Contaminación del Aire, con el objetivo de determinar la efectividad de las medidas de prevención y control implementadas, y planear recomendaciones que permitan mejorar la gestión de la calidad del aire en el país. Producto 4: informe de resultados de la evaluación y segunda entrega de documentación de las bases de datos. DNP. Bogotá. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2017b). CTTO DNP-605 de 2016 Realizar una evaluación de resultados de la Política de Prevención y Control de la Contaminación del Aire, con el objetivo de determinar la efectividad de las medidas de prevención y control implementadas, y planear recomendaciones que permitan mejorar la gestión de la calidad del aire en el país. Producto 5: Plan de Transferencia e Implementación de Recomendaciones – PTIR. DNP. Bogotá. Colombia.
- Economía Urbana e Infométrika (2015a). Evaluación institucional y de resultados del programa “Planes Departamentales para el manejo Empresarial de los Servicios de Agua y Saneamiento” específicamente los sectores de acueducto y alcantarillado, que permita identificar la capacidad institucional de los actores para cumplir con los objetivos del PDA y también el grado de cumplimiento de los objetivos y metas del programa, en el periodo 2008-2013. Producto 4: Informe de Resultados de la Evaluación, Análisis Costo Beneficio y tercera entrega de documentación bases de datos. DNP. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2015b). Evaluación institucional y de resultados del programa “Planes Departamentales para el manejo Empresarial de los Servicios de Agua y Saneamiento. Producto 5: Plan de Transferencia e Implementación de Recomendaciones. DNP. Bogotá. Colombia.
- Evans, P.B. et al. (1985). Bringing the State Back. Cambridge University Press, Cambridge.
- Fischer, Frank (2009). Democracy and Expertise. Oxford University Press. Oxford.
- G/Exponencial (2015a). Producto 4 Informe de Resultados de la Evaluación. Evaluación institucional y de resultados de la Política Nacional de Seguridad Alimentaria y Nutricional –PSAN-. DNP. Bogotá. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145

- _____ (2015b). Evaluación institucional y de resultados de la Política Nacional de Seguridad Alimentaria y Nutricional –PSAN- (Ficha técnica). DNP. Bogotá. Colombia. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2015c). Producto 5 Plan de Transferencia e Implementación de Recomendaciones –PTIR- Evaluación institucional y de resultados de la Política Nacional de Seguridad Alimentaria y Nutricional –PSAN-. DNP. Bogotá. Colombia.
- _____ (2016a). Evaluación de operaciones que permita identificar las fortalezas y debilidades de las actividades que se ejecutan en el proceso de investigación e implementación o prórroga de una medida de defensa comercial (Salvaguardia o Derecho Antidumping), con el fin de generar recomendaciones conducentes a mejorar la calidad y pertinencia del mismo (Ficha técnica). Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2016b). Evaluación de operaciones que permita identificar las fortalezas y debilidades de las actividades que se ejecutan en el proceso de investigación e implementación o prórroga de una medida de defensa comercial (Salvaguardia o Derecho Antidumping), con el fin de generar recomendaciones conducentes a mejorar la calidad y pertinencia del mismo. Producto 3 Informe de Resultados de la Evaluación. Contrato DNP-494-2016. Recuperado de: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- _____ (2016c). Evaluación de operaciones que permita identificar las fortalezas y debilidades de las actividades que se ejecutan en el proceso de investigación e implementación o prórroga de una medida de defensa comercial (Salvaguardia o Derecho Antidumping), con el fin de generar recomendaciones conducentes a mejorar la calidad y pertinencia del mismo. Producto 4 Plan de Transferencia e Implementación de Recomendaciones - PTIR. Contrato DNP-494-2016.
- Garson, G. David (1992). De la ciencia de las políticas al análisis de políticas: veinticinco años de progreso. En: Aguilar, Luis. El Estudio de las Políticas Públicas. Grupo Editorial Miguel Angel Porrua. México.
- Gerencial Ltda. (2013a). Evaluación intermedia de la operación y los resultados del Sistema Nacional de Evaluación de Gestión y Resultados –SINERGIA-. DNP. Bogotá. Colombia.
- _____ (2013b). Ficha técnica de: Evaluación intermedia de la operación y los resultados del Sistema Nacional de Evaluación de Gestión y Resultados –SINERGIA-. DNP. Bogotá. Colombia. Recuperada en: https://sinergiapp.dnp.gov.co /#Evaluaciones/EvalFin/145
- Gobierno de Colombia (2013). Plan Nacional de Seguridad Alimentaria y Nutricional 2012-2019. Recuperado de: https://www.icbf.gov.co/sites/default /files/pnsan.pdf
- González Peña, José Danilo (2016). Por qué se usa la información de seguimiento y evaluación (SyE) de políticas públicas. Centro de Investigación y Docencia Económicas (CIDE), México.

- Grupo de Planeación-DNP (2010). Contratación de la evaluación. Código: SE-EP-CE. DNP. Bogotá. Colombia.
- _____ (2013a). Diseño de la evaluación. Código: SE-EP-DE. DNP. Bogotá. Colombia.
- _____(2013b). Ejecución de la Evaluación. Código: SE_EP_EE. Versión 4. Grupo de Planeación del DNP. Bogotá. Colombia
- _____(2013c). Socialización de los resultados de la evaluación. Código: SE_EP_UR. Versión 2. DNP. Bogotá. Colombia.
- _____(2016). Diseño de la evaluación Código: SE_EP_DE, versión 4. DNP. Bogotá. Colombia.
- Gschwend, Thomas y Schimmelfennig, Frank (2007). Research Design in Political Science. Palgrave Macmillan. London.
- Guba, Egon (1990). The Paradigm Dialog. Sage Publications. London.
- Guba, E. y Lincoln, Y. (1989). Fourth Generation Evaluation. Sage Publications. Londres.
- Habermas, Jürgen (1998). Facticidad y Validez. Trotta. Madrid. España.
- Hall, Peter (1986). Governing the Economy. The Politics of State Intervention in Britain and France. Polity. Oxford
- _____ (1993). Policy Paradigms, Social Learning and the State: The Case of Economic Policymaking in Britain. Comparative Politics. 3.
- Hatry, Harry (2002). Performance Measurement, Fashions and Fallacies. Public Performance & Management Review, 25
- Hofstetter, Carolyn. H. y Alkin, Marvin (2003). Evaluation use revisited. En: Kellaghan; Stufflebeam y Wingate. International handbook of education evaluation. Kulwer. Boston. EEUU.
- Holland, Jhon; Holyoak, Keith; Nisbett, Richard y Thagard, Paul (1986). Induction: Process of inference, Learning and Discovery. Cambridge: MIT: Press.
- Johansson, Tobias y Siverbo, Sven (2009). Explaining the Utilization of Relative Performance Evaluation in Local Government: A Multi-Theoretical Study Using Dara for Sweden. Financial Accountability & Management, 25(2). EEUU.
- Julnes, Patria y Holzer, Marc (2001). Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation.
- King, Gary; Keohane, Robert; Verba, Sidney (2000). El diseño de la investigación social, la inferencia científica en los estudios cualitativos. Alianza Editorial. Madrid.
- Kingdon Jhon (1993). “How do Issues get on Public Policy Agenda?”, en Wilson W. J. (ed.) Sociology and the Public Agenda, Sage, Newbury Park. EEUU.
- Knoepfel, Peter, Larrue, Corinne, Varone, Frédéric (2006). Analyse et pilotage des politiques publiques. Verlag Rüegger. Zürich.
- Kroll, Alexander (2015). Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. Public Performance & Management Review 38(3):459-486.

- Kusek, Jody y Rist, Ray (2005). Diez pasos hacia un sistema de seguimiento y evaluación basado en resultados. Banco Mundial. Mayol ediciones. Bogotá.
- Lasswell, Harold (1951). La orientación hacia las políticas. En: Aguilar V., L.F., (Ed.), El estudio de las políticas públicas, Miguel Ángel Porrúa, México.
- Lindblom, Charles (1990). Inquiry and Change: The Troubled Attempt to Understand and Shape Society. Yale University Press. EEUU.
- Majone, Giandomenico. (1997). Evidencia, argumentación y persuasión en la formulación de políticas, FCE, México.
- March, James G., Olsen, Johan P. (1997) El redescubrimiento de las instituciones. La base organizativa de las políticas. Fondo de Cultura Económica. Mexico.
- Ministerio de Ambiente, Vivienda y Desarrollo Territorial (2010). Política de Prevención y Control de la Contaminación del Aire. Ministerio de Ambiente, Vivienda y Desarrollo Territorial. Bogotá. Colombia. Recuperado de: http://www.minambiente.gov.co/images/AsuntosambientalesySectorialyUrbana/pdf/Polit%C3%ACcas_de_la_Direcci%C3%B3n/Pol%C3%ADtica_de_Prevenci%C3%B3n_y_Control_de_la_Contaminaci%C3%B3n_del_Aire.pdf
- Ministerio de Vivienda, Ciudad y Territorio (2013). Programa de Conexiones Domiciliarias e Intradomiciliarias de Acueducto y Alcantarillado. Presentación en Powerpoint. MVCT. Colombia.
- Morra, Linda y Rist, Ray (2009). The Road to Results. Designing and conducting effective development evaluations. The World Bank. Washington, DC. EEUU.
- Moynihan, Donald y Pandey, Sanjay (2010). The Big Question for Performance Management: Why Do Managers Use Performance Information? Journal of Public Administration Research and Theory, 20.
- Moynihan, Donald; Pandey, Sanjay y Wright, Bradley (2012). Setting the Table: How Transformational Leadership Fosters Performance Information Use. Journal of Public Administration Research and Theory, 22.
- Muller, Pierre. (2002). Las políticas Públicas. Universidad Externado de Colombia. Bogotá.
- North, Douglas C. (1990) Institutions, institutional Change and Economic Performance, Cambridge University Press, Cambridge.
- OCDE (2005). Modernising government: the way forward. OCDE. París. Francia.
- _____ (2010). Estándares de Calidad para la Evaluación del Desarrollo. Comité de Ayuda al Desarrollo –CAD-. OCDE. París. Francia.
- ONU (1998). Informe del Representante del Secretario General Sr. Francis M. Deng, presentado con arreglo a la resolución 1997/39 de la Comisión de Derechos Humanos. Principios Rectores de los desplazamientos internos. Consejo Económico y Social. Naciones Unidas. Recuperado de: https://www.acnur.org/fileadmin/Documentos/BDL/2001/0022.pdf
- Osborne, David y Gaebler, Ted (1994). La Reinvención del Gobierno. La influencia del Espíritu Empresarial en el Sector Público. Ediciones Paidós. Barcelona.

- Ostrom, Elinor (1990). Governing the Commons. The Evolution of Institutions for Collective Action. Cambridge University Press. Cambridge.
- _____ (2007). Institutional Rational Choice: An Assessment of the Institutional Analysis and Development Framework. En: Sabatier, Paul. Theories of the Policy Process. Westview Press, Boulder, Colorado, EEUU.
- _____ (2015). Comprender la diversidad institucional. Fondo de Cultura Económica. México.
- Parsons, Wayne, 2007. Políticas públicas, FLACSO, México.
- Patton, Michael (2002). Qualitative Research & Evaluation Methods. 3th edition. Sage Publications. London.
- Patton, Michael (2012). Essentials of Utilization-Focused Evaluation. Sage Publications. Thousand Oaks. EEUU.
- Patton, Michael (2013). Utilization-Focused Evaluation (U-FE) Checklist: https://wmich.edu/sites/default/files/attachments/u350/2014/UFE_checklist_2013.pdf
- Peck, Laura y Gorzalski, Lindsey (2009). An Evaluation Use Framework and Empirical Assessment. Journal of Multidisciplinay Evaluation, 6(12).
- Peters, Guy (2003). El nuevo institucionalismo. Gedisa. Barcelona. España.
- Perez, Anibal (2007). El Método Comparativo: Fundamentos y Desarrollos Recientes, Universidad de Pittburgh. Pittsburgh. EEUU.
- Powell, Walter W., DiMaggio, Paul J. (1999). El nuevo institucionalismo en el análisis organizacional. FCE. México.
- Ragin, Charles (2017) User's guide to Fuzzy-Set/Qualitative Comparative Analysis. University of California. Se encuentra en: http://www.socsci.uci.edu/~cragin/fsQCA/software.shtml
- Ramírez, María (2009a). Las reformas del Estado y la administración pública en América Latina y los intentos de aplicación del New Public Management. En: Estudios Políticos, 34, Instituto de Estudios Políticos, Universidad de Antioquia. Medellín.
- _____ (2009b). Transferencia de Políticas y Reforma de la Administración Pública en Colombia. Universidad Autónoma de Madrid. Madrid.
- Rihoux, Benoit y Ragin, Charles (2009). Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques. SAGE. Thousand Oaks. EEUU.
- Roe, Emery (1994). Narrative Policy Analysis. Duke University Press. London.
- Roth, André-Noël (2007) Una introducción a la evaluación de las políticas públicas. En: Vargas A. et al. Fundamentos de Políticas Públicas. Universidad Nacional de Colombia. Bogotá. Colombia.
- _____ (2015). Neo-institucionalismo y transformación democrática del Estado. En: El rol del Estado: contribuciones al debate. Friedrich Ebert Stiftung.
- _____ (2016). El análisis de las políticas públicas: situación y retos para América Latina. Conferencia magistral presentada como requisito para ascenso a Profesor Titular de la Universidad Nacional de Colombia. Facultad de

Derecho, Ciencias Políticas y Sociales, Auditorio Camilo Torres, 29 de noviembre del 2016. Bogotá. Colombia.

- Sabatier, Paul (2007). Theories of the Policy Process. Westview Press. Boulder. EEUU.
- Saliterer, Iris y Korac, Sanja (2013). Performance information use by politicians and public managers for internal control and external accountability purposes. Critical Perspectives on Accounting, 24. EEUU.
- Sandison, Peta (2006). The Utilization of evaluations. En ALNAP, Review of Humanitarian Action in 2005: evaluation utilization. United Kingdom: ALNAP.
- Schmidt, Vivien (2008). "Discursive Institutionalism: The Explanatory Power of Ideas and Discourse". Annual review of Political Science, 11.
- Schwand, Thomas (1992) Constructing appropiateand useful metaevaluative frameworks. Evaluating and Program Planning. Vol 15, pp. 95 -100. USA.
- Schwandt, T. A., & Halpern, E. S. (Eds.) (1988). Linking Auditing and Meta-Evaluation: Enhancing Quality in Applied Research. (Applied Social Research Methods; Vol. 11). Sage Publications.
- Scott, W. R., Meyer, J. W. (1994). Institutional Environments and Organizations: Structural Complexity and Individualism, sage Publications, Thousanks Oaks. EEUU.
- Scriven, M. (2010) Evaluating Evaluations: a meta-evaluation checklist. 9th edition, Claremont Graduate University.
- Scriven, M. (2015) Key evaluation checklist (KEC)- Claremont Graduate University & The Evaluation Center, Western Michigan University
- Solarte, Leonardo (2004). Las Evaluaciones de Políticas Públicas en el Estado Liberal, Universidad del Valle. Cali.
- Steinmo et al., (1992). Structuring Politics. Historical Institutionalism in Comparative Analysis, Cambridge University Press, New York. EEUU
- Stufflebeam, D. (1999) Program Evaluations Metaevaluation Checklist (Bases on the Program Evaluation Standards. Se encuentra en: www.wmich.edu/evalctr/checklists
- Stufflebeam, D. (2001) The metaevaluation imperative. American Journal of Evaluation, Vol 22 (2). 183 - 209. Michigan.
- Stufflebeam, D; Goodyear, L; Marquart, J; Johnson E. (2005) Guiding Principles checklist for evaluating evaluations
- Stone, Deborah (1988). Policy Paradox and Political Reason. W. W. Norton & Company. New York. EEUU.
- Stone, Deborah (1989). "Causal Stories and the Formation of Policy Agendas". En: Political Science Quarterly, 104 (2).
- Taylor, Jeannette (2011). Factors Influencing the Use of Performance Information for Decision Making in Australian State Agencies. Public Administration, 89 (4).
- Torres L; Pina V y Yetano A. (2011). Performance measurement in Spanish local governments. A cross-case comparison study. Public Admin. 89 (1).
- UARIV (2015). Mapa conceptual de la Política Pública para la Prevención, Protección, Atención, Asistencia y Reparación Integral a las Víctimas.

Recuperado de: https://www.unidadvictimas.gov.co/es/snariv/mapa-conceptual-de-política-pública/13363

- _____ (2017a). Protocolo para el acompañamiento a retornos y reubicaciones en el marco de la reparación integral a víctimas del desplazamiento. Recuperado de: https://www.unidadvictimas.gov.co/es/protocolo-acompanamiento-retornos-y-reubicaciones-en-el-marco-de-la-reparacion-integral-victimas-del
- _____ (2017b). Procedimiento para la Gestión de la Implementación y Seguimiento de los PIRC. Sistema Integrado de Gestión. UARIV. Bogotá. Colombia.
- _____ (2018a). Manual Operativo Ruta de Atención, Asistencia y Reparación Integral a las Víctimas. Proceso: Servicio al Ciudadano. UARIV. Bogotá. Colombia.
- _____ (2018b). Manual metodológico para la formulación de PIRC bajo la estructura de marco lógico. Procedimiento: Diseño, formulación y aprobación del Plan Integral de Reparación Colectiva para sujetos no étnicos. UAIRV. Bogotá. Colombia.
- United Nations Evaluation Group –UNEG- (2005). “Standards for Evaluation in the UN System”. ONU.
- Uusikylä, Petri y Virtanen, Petri (2000) MetaEvaluation as a Tool for Learning. A case study of the European structural Fun Evaluations in Finland. En Evaluation, Vol 6 (I) pp. 50-65. SAGE Publications, Londres.
- Van Dooren W. (2008). Nothing new under the sun? Change and continuity in twentieth century performance movements. En: van Dooren W, van de Walle ST, editors. Performance information in the public sector. How it is used. Palgrave Macmillan. Houndmills. Reino Unido.
- Van de Walle, S. y Van Dooren, W. (2010). How is Information Used to Improve Performance in the Públic Sector? Exploring the Dynamics of Performance Information. En K. Walshe; G. Harvey, y P. Jas (eds.) Connecting Knowledge and Performance in Public Services: From Knowing to Doing. Oxford: Oxford University Press.
- Waever, R. K. y Rockman, B. A. (1993). Do Institutions Matter? Government Capabilities in the United States and Abroad, The Brooking Institution, Washington
- Weiss, Carol H. (1975). Investigación evaluativa. Métodos para determinar la eficiencia de los programas de acción. Quinta reimpresión (1999) Trillas. México.
- Williamson, O. (1985) The Economic institutions of Capitalism, Free Press, New York. EEUU.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011). The program evaluation standards: A guide for evaluators and evaluation users (3rd ed.). Sage. Thousand Oaks, EEUU.
- Zittoun, Philippe (2014). The Political Process of Policymaking. A Pragmatic Approach to Public Policy, France: Palgrave McMillan, University of Lyon.

Normatividad:

- Acuerdo Antidumping de la OMC de 1994
- Acuerdo General sobre Aranceles Aduaneros y Comercio (GATT) de 1994
- CONPES 3463 de 2007
- CONPES 113 de 2008
- CONPES 3943 de 2018
- Decreto 2055 de 2009
- Decreto 4633 de 2011
- Decreto 4634 de 2011
- Decreto 4635 de 2011
- Decreto 4800 de 2011
- Decreto 1290 de 2014
- Decreto 1082 de 2015
- Decreto 1084 de 2015
- Decreto 1750 de 2015
- Decreto 2189 de 2017
- Decreto 2439 de 2018
- Ley 337 de 1997
- Ley 488 de 1998
- Ley 152 de 1994
- Ley 1448 de 2011
- Resolución 2043 de 2012.
- Resolución 3789 de 2014 del DNP
- Resolución 4219 de 2014 del DNP
- Resolución 4220 de 2014 del DNP
- Resolución 4287 de 2014 del DNP
- Resolución 3903 de 2015 del DNP
- Resolución 4386 de 2016 del DNP
- Resolución 2254 de 2017
- Resolución 3143 de 2018 de la UARIV
- Sentencia de tutela T025 de 2004

Respuestas derechos de petición:

- DNP, respuesta derecho de petición N° 20183600182751
- respuesta a derecho de petición del DNP N° 20193600428271