

¿Cómo incide el método de las evaluaciones en su uso?

José Danilo González Peña

José Danilo González Peña

Economista de la Universidad Santo Tomás, título de Maestría en Análisis de Problemas Políticos, Económicos e Internacionales Contemporáneos de la Universidad Externado de Colombia y Doctor en Políticas Públicas del Centro de Investigación y Docencia Económicas (CIDE) en México. Trabajó para el Departamento Nacional de Planeación de Colombia en las direcciones de Seguimiento y Evaluación de Políticas Públicas y de Inversiones y Finanzas Públicas. Fue docente en universidades como la Nacional, la Central y Externado de Colombia. Desde 2018 es profesor asistente en la Escuela de Gobierno Alberto Lleras Camargo de la Universidad de los Andes.

Las comunicaciones con el autor pueden dirigirse a:
E-mail: jd.gonzalezp@uniandes.edu.co

¿Cómo incide el método de las evaluaciones en su uso?

Las crecientes necesidades de los ciudadanos en diferentes aspectos de la vida diaria, junto a las enormes restricciones presupuestales para cubrir las adecuadamente, obligan a los gerentes públicos a llevar a cabo políticas mucho más eficientes y efectivas. Una de las principales funciones de las evaluaciones es servir de insumo en la toma de decisiones para mejorar las políticas públicas y en procesos relacionados con la rendición de cuentas del gobierno. No obstante, en Latinoamérica, y específicamente en Colombia, existen escasos análisis sobre el uso de la información de las evaluaciones de programas y políticas públicas. Para avanzar en el cierre de esta brecha se realizó un estudio descriptivo y un análisis multivariado en Colombia que analiza los usos potenciales de la información evaluativa y la incidencia del método utilizado en los diferentes tipos de uso. Para efectuar análisis se utilizó información de 251 evaluaciones provenientes de los sectores educación, salud, inclusión social, agricultura, transporte e infraestructura, y vivienda a nivel nacional para el periodo comprendido entre 2008 y 2019. Los principales resultados de este análisis son: i) la mayoría de las evaluaciones muestran un uso potencial de tipo instrumental, es decir, sus hallazgos y recomendaciones pretenden ser empleadas para tomar decisiones directas e inmediatas sobre el programa evaluado; ii) 23,5% de las evaluaciones presenta un uso potencial de tipo conceptual, es decir, fueron diseñadas para brindar información para ejercicios de rendición de cuentas o mejorar el conocimiento sobre la forma como funciona u opera el programa analizado; iii) el uso instrumental tiene una relación estadísticamente significativa y positiva con los métodos cualitativos, la contratación pública y la evaluación concomitante; y iv) el uso conceptual tiene una relación estadísticamente significativa y negativa con los métodos cualitativos y con los sectores agricultura, transporte y vivienda.

Palabras clave: Evaluación; Modelos de Evaluación; Evaluación de Políticas; Evaluación de Programas; Métodos; Toma de Decisiones; Colombia

How Does the Method of Evaluations Affect their Use?

The growing needs of citizens in different aspects of daily life, together with the enormous budgetary restrictions to adequately cover them, force public managers to carry out much more efficient and effective policies. One of the main functions of evaluations is to serve as an input in decision-making to improve public policies and in processes related to government accountability. However, in Latin America, and specifically in Colombia, there is little analysis on the use of information from evaluations of public

Recibido: 23-07-2022 y 15-11-2022 (segunda versión). Aceptado: 24-11-2022.

programs and policies. To make progress in closing this gap, a descriptive study and a multivariate analysis were carried out in Colombia to analyze the potential uses of evaluative information and the incidence of the method used in the different types of use. The analysis used information from 251 evaluations from the education, health, social inclusion, agriculture, transportation and infrastructure, and housing sectors at the national level for the period 2008-2019. The main results of this analysis are: i) most of the evaluations show a potential instrumental use, i.e. their findings and recommendations are intended to be used to make direct and immediate decisions about the evaluated program; ii) 23.5% of the evaluations show a potential conceptual use, i.e. they were designed to provide information for accountability exercises or to improve knowledge about how the analyzed program works or operates; iii) instrumental use has a statistically significant and positive relationship with qualitative methods, public procurement and concomitant evaluation; and iv) conceptual use has a statistically significant and negative relationship with qualitative methods and with the agriculture, transportation and housing sectors.

Key words: Evaluation; Evaluation Models; Policies Evaluation; Programs Evaluation; Methods; Decisions Making; Colombia

Introducción

Las crecientes necesidades de los ciudadanos en diferentes aspectos de la vida diaria, como la salud, la educación, el transporte, la vivienda, etc., junto a las enormes restricciones presupuestarias para cubrirlas adecuadamente, obligan a los gerentes públicos a llevar a cabo políticas mucho más eficientes y efectivas. Contar con información oportuna y veraz para la toma de decisiones, como las evaluaciones, garantiza que los programas mejoren continuamente sus servicios (Stufflebeam, 1983; Alkin y Taut, 2002) y se convierte en un elemento fundamental en contextos de alta incertidumbre. Asimismo, las evaluaciones pueden contribuir al fortalecimiento de la institucionalidad democrática y a la legitimidad del Estado, en la medida en que promuevan los procesos de transparencia entre el gobierno y la ciudadanía y logren establecer niveles de responsabilidad claros entre los funcionarios públicos y su obligación de rendir cuentas (Cunill y Ospina, 2003).

No obstante, a pesar de los esfuerzos normativos, institucionales y presupuestarios que muchos gobiernos han desplegado para la implementación y uso de sistemas de seguimiento y evaluación en las últimas dos décadas (Taylor, 2011a; Wholey ...[et al], 2010; Cunill y Ospina, 2003; Newcomer, 1997), bajo el paraguas de los enfoques de la Nueva Gestión Pública (Lapiente y Van de Walle, 2020; Spekle

Para algunos investigadores la causa de este limitado uso de la evidencia es que las evaluaciones no generan información relevante y oportuna para tomar decisiones.

y Verbeeten, 2009) y la Gestión Pública Orientada a Resultados, la evidencia indica que esta información no ha permeado la toma de decisiones de los gerentes públicos (Mikkelsen ...[et al], 2021; Kumar ...[et al], 2020; Newman ...[et al], 2017; Hojlund, 2014; Askim, 2007; Heinrich y Marschke, 2010; Moynihan y Pandey, 2010; Weiss, 1982). Por ejemplo, algunas estimaciones en EE. UU. muestran que bajo las dos administraciones del presidente Obama, solo 1% de la asignación del gasto público se basó en evidencias provenientes de evaluaciones (Bridgeland y Orszag, 2013; OCDE, 2020). Además, las evaluaciones también pueden ser utilizadas para alcanzar objetivos personales y políticos que estarían relacionados poco o nada con el fortalecimiento de la democracia y el logro de un mayor bienestar de la sociedad (Shaxson y Boaz, 2020; Leviton y Hughes, 1981; Oszlak, 2013; Henry, 2000; Hyyryläinen y Viinamäki, 2008).

La preocupación sobre el mejor uso de la evidencia y la información disponible por parte de los gerentes públicos no es un tema reciente. Esta inquietud ha sido un tema recurrente de análisis en el campo de la administración y las políticas públicas por décadas, como lo resaltaron Arellano y Hernández (2014). No obstante, para académicos como Shaxson (2019) todavía hace falta mayor investigación empírica y conceptual sobre este importante tema.

Para algunos investigadores la causa de este limitado uso de la evidencia es que las evaluaciones no generan información relevante y oportuna para tomar decisiones (Newcomer, 1997; Moynihan e Ingraham, 2001). En este sentido, Pawson ...[et al] (2011) sostienen que la evidencia muy pocas veces ofrece la certeza y la seguridad que demandan las decisiones políticas, sino que más bien se puede considerar como un proceso acumulativo en el que los datos persiguen, pero nunca captan del todo los problemas de la política pública. En este sentido, si se quiere incrementar de forma importante el uso de la evidencia, se debe incluir y trabajar de forma colaborativa con los actores clave (*stakeholders*) que se relacionan con el programa a evaluar (Owen y Lambert, 1998) porque, en definitiva, la función del evaluador se debe enfocar en servir de facilitador que brinda a los clientes el conocimiento y las herramientas para una autoevaluación y una rendición de cuentas permanentes (Fetterman, 1996).

Otros autores opinan que los incentivos que brindan las organizaciones no son suficientes para que los gerentes utilicen dicha información (Moynihan y Hawes, 2012; Taylor, 2011a). También algunos académicos consideran que existen atributos individuales como la

actitud, la motivación, los hábitos y las costumbres, entre otros, que no facilitan el uso de la información para la toma de decisiones en las organizaciones (Kroll, 2012; Van de Walle y Van Dooren, 2009).

Finalmente, otros consideran que los procesos de generación y uso de la evidencia para tomar decisiones no se desarrollan en un vacío (Cárdenas ...[et al], 2022), sino que se realizan en un contexto político-social que frecuentemente es ignorado en el análisis de las políticas públicas (Weiss, 1982; Moynihan y Pandey, 2010; Moynihan y Hawes, 2010; Saliterer y Korac, 2013; Kroll, 2012; Balthasar, 2009; Appleton-Dyer ...[et al], 2012; Parkhurst, 2016; De Lancer Julnes y Holzer, 2001, entre otros).

La comprensión adecuada de este ámbito permitiría contrastar parcialmente lo que se observa entre la conceptualización de un modelo de gestión, que se orienta de manera especial hacia la evaluación de intervenciones gubernamentales, y aquellos factores que inciden en el limitado uso de información (Cárdenas ...[et al], 2022; Ledermann, 2012; Moynihan, 2005; Aiello ...[et al], 2021; Cairney, 2019; Choi y Woo, 2021; French, 2018; McDonnell y Weatherford, 2020; Pallett, 2020).

Si bien existe abundante literatura académica que aborda la inquietud sobre el uso limitado de la información evaluativa para tomar decisiones (Preskill, 2004; Burr, 2009; De Lancer Julnes, 2006; Moynihan y Pandey, 2010; Moynihan y Hawes, 2010; Johansson y Siverbo, 2009; Newcomer, 1997), no se observa un gran énfasis hacia otros tipos de uso. Asimismo, gran parte de los estudios empíricos hacen referencia a organizaciones que se encuentran en países industrializados, pero no a países latinoamericanos. Y estos análisis se enfocan más en tipos de uso efectivo de la información, pero no en usos esperados o potenciales.

Tener un conocimiento limitado del uso de la información de las evaluaciones en el contexto latinoamericano (Cárdenas ...[et al], 2022; Luna, 2016; Zaltsman, 2010; Cardozo, 2006, 2013a y 2013b; González, 2016) puede tener varias implicaciones: i) no es posible saber si las evaluaciones están cumpliendo adecuadamente con los objetivos propuestos, lo que puede llevar a un despilfarro de recursos; ii) es posible que no se evidencie la importancia de las evaluaciones dentro del ciclo de políticas públicas y terminen perdiendo relevancia en la administración pública; iii) dificulta tener un conocimiento veraz y oportuno sobre cuáles pueden ser los posibles inconvenientes que impidan consolidar el uso de dicha información; y iv) puede brindar

incentivos perversos para que solamente se terminen usando aquellas evaluaciones que muestran resultados positivos.

En este contexto, se considera relevante responder las siguientes preguntas para Colombia: ¿Para qué se utiliza la información de las evaluaciones (tipos de uso)? ¿Cómo incide el método de las evaluaciones de políticas públicas en su uso? Actualmente no existen trabajos académicos en el país que exploren en detalle estas preguntas. Para avanzar hacia posibles respuestas es relevante realizar un mapeo de las principales evaluaciones que existen y analizar el tipo de uso potencial de estos estudios, a partir de los propósitos establecidos en estos documentos. Este es el objetivo central de este documento que está organizado en cuatro secciones. En la primera se realiza una breve descripción de la principal literatura relacionada con el uso de la información de las evaluaciones. En la segunda sección se explica la metodología utilizada para la recolección y análisis de la información. La tercera parte incluye los principales resultados del análisis. Finalmente se esbozan las conclusiones del ejercicio.

1. Literatura y antecedentes

En este documento se entiende el concepto evaluación como una valoración sistemática y objetiva de un proyecto, programa o política en relación con su diseño, ejecución o resultados. Su finalidad es determinar la pertinencia y el logro de objetivos, la eficiencia del desarrollo, su efectividad, impacto y sostenibilidad (OCDE, 2002).

De acuerdo con Rossi ...[et al] (2004) y Chelimsky (2006), la información proveniente de las evaluaciones tiene los siguientes propósitos:

- Mejoramiento de los programas. Los hallazgos de las evaluaciones pueden generar información para mejorar el programa. Esta información le puede servir a los diferentes involucrados en el proceso de evaluación para responder preguntas relacionadas con la necesidad del programa, el diseño y el concepto del programa, su implementación, su impacto o su eficiencia.

- Rendición de cuentas (*accountability*). Sirve para determinar si se cumplieron las expectativas programadas. Asimismo, esta información puede influir sobre diferentes decisiones como la continuación del programa, la asignación de recursos, reestructuraciones o acciones legales (Rossi ...[et al], 2004).

- Generar nuevo conocimiento. Esto significa responder preguntas de investigación sobre aspectos desconocidos de un programa o una política pública. Puede ser utilizada para temas de investigación

Desde los años ochenta, varios académicos se dieron a la tarea de empezar a estudiar el término de uso de evaluaciones en la toma de decisiones.

y para conocer, por ejemplo, cómo una determinada intervención puede ayudar a mejorar la implementación de un programa.

- Mejorar la capacidad de las agencias. Es decir, utilizar la información evaluativa para que la agencia pueda resolver problemas relacionados con los programas que coordina, desarrollar bases de datos, habilidades analíticas y prácticas gerenciales, entre otros.

- Ardides (artimañas) políticos o relaciones públicas. A veces la información que proporcionan las evaluaciones tiene como propósito brindar un contexto político para una decisión que ha sido tomada con anterioridad. También puede utilizarse como táctica dilatoria para apaciguar las críticas o aplazar decisiones difíciles (Rossi ...[et al], 2004).

Halachmi (2002), resume las dos razones principales para usar la información evaluativa: para fortalecer los procesos de rendición de cuentas (*accountability*) y para mejorar el desempeño (de las organizaciones y de las políticas públicas). Ongaro y Andreotti (2022) relacionan el uso de información en los procesos de toma de decisiones en las organizaciones con la reducción de la incertidumbre.

Por otra parte, para clasificar las evaluaciones se emplean dos categorías descritas por Ordóñez-Matamoros ...[et al] (2018), esto es según el tiempo y el objeto. De acuerdo con el tiempo, las evaluaciones se clasifican así: i) *ex ante*, es decir, con una orientación prospectiva o que efectúan un ejercicio diagnóstico sobre los posibles efectos; ii) *concomitante*, aquellas evaluaciones que se desarrollan durante la ejecución de un programa para corregir su curso; y iii) *ex post*, comprenden las evaluaciones con una orientación retrospectiva y busca analizar los efectos reales de la intervención.

De acuerdo con su objeto, las evaluaciones se pueden categorizar en: i) insumos, aquellas que implican un análisis sobre los recursos, capacidades y personas empleadas; ii) procesos, las que realizan un análisis sobre las actividades de transformación de insumos para generar productos y resultados; iii) resultados, es decir, que efectúan un análisis sobre los productos generados con el programa; y iv) impactos, las evaluaciones que realizan un análisis sobre los efectos no intencionados positivos y negativos (Ordóñez-Matamoros ...[et al], 2018).

1.1 Concepto de uso y tipos de uso de la información evaluativa

El concepto de uso de información de evaluaciones ha ido cambiando a lo largo del tiempo. Desde los años ochenta, varios académicos se dieron a la tarea de empezar a estudiar el término de *uso de evaluaciones* en la toma de decisiones. De estas primeras investigaciones,

sigue siendo relevante en la actualidad y constituye la base conceptual de este documento la clasificación propuesta por Leviton y Hughes (1981) sobre los tipos de uso de la información de los documentos evaluativos. Según esta definición la información puede usarse de forma *instrumental, conceptual y simbólica o persuasiva* (Alkin y Taut, 2002; Fleischer and Christie, 2009; Ledermann, 2012). Aunque existen otras clasificaciones de diversos autores sobre los usos de la información evaluativa (Moynihan, 2009; Van Dooren ...[et al], 2010; Kroll, 2012; Sandison, 2006; entre otros), se puede considerar la clasificación de Leviton y Hughes como la más difundida en el ámbito académico (Cárdenas ...[et al], 2022; Balthasar, 2009; Cousins ...[et al], 2014; Hyyryläinen y Viinamäki, 2008; Lammert ...[et al], 2017).

El uso instrumental hace referencia a la aplicación directa e inmediata de los resultados de la evaluación a decisiones específicas tales como financiar o no un programa, decisiones sobre su gestión o funcionamiento, o como una revisión para mejorar (Cousins ...[et al], 2014). Por tanto, este es un tipo de uso esperado conforme a un modelo racional-burocrático en el que el conocimiento y la información generados por investigadores o analistas se incorporan de forma directa y cíclica en los procesos decisorios (Cárdenas ...[et al], 2022). Este uso se incrementa cuando los mismos responsables de las mediciones son quienes cubren los costos y la evaluación está orientada a los requerimientos de los usuarios potenciales (Balthasar, 2009).

Para Peck y Gorzalski (2009) si se desea incrementar el uso instrumental, las recomendaciones de las evaluaciones deben estructurarse mucho mejor: se debe tener en cuenta el contexto del programa y la capacidad de las agencias, deben indicar con claridad quién debe implementarlas (UNEG, 2016) e incrementar el contacto entre los evaluadores y las agencias evaluadas.

El uso conceptual hace referencia a aquellos casos en donde la información es utilizada para educar, cambiar o profundizar la perspectiva y el conocimiento que los tomadores de decisiones tienen sobre un asunto, programa o política, así como para dar a conocer los efectos intencionados y no intencionados del programa (Cousins ...[et al], 2014; Boyer y Langbein, 1991). Este también puede cumplir el rol de uso intermedio, como un aprendizaje previo para un uso instrumental posterior (González, 2016). Además, se relaciona con la rendición de cuentas al presentarse como un proceso de control, donde las personas responden y son responsables por el funcionamiento del programa (Luna, 2016).

Los gerentes públicos toman decisiones utilizando una racionalidad intencional, pero sus capacidades cognitivas están acotadas y la información con la que cuentan es incompleta.

El uso simbólico o persuasivo hace referencia al uso de información para justificar una posición o decisión política adoptada con anterioridad, o por razones distintas a los resultados de la evaluación. También en casos donde se busca reafirmar el valor del programa o cuando se busca dar cumplimiento a los mandatos del financiador del programa (Cousins ...[et al], 2014). Algunos autores consideran esta categoría de uso como un mal uso de la evidencia (Shulha y Cousins, 1997), debido a que puede desplazar la relevancia que adquiere usar la información para la concreción específica de resultados de una intervención, de manera que el uso de información resulte directamente en un mejoramiento social (Henry, 2000).

Más recientemente se han propuesto aproximaciones diferentes que sugieren otras clasificaciones de usos, como la de Rickinson ...[et al] (2017) quienes identificaron hasta doce usos posibles de la evidencia y la información en organizaciones gubernamentales, agrupados alrededor a tres grandes objetivos en el proceso de las políticas públicas: i) para la definición y sensibilización sobre la existencia de problemas públicos; ii) para el diseño y la selección de intervenciones; y iii) para la creación de una narrativa que permita garantizar la continuidad de determinado programa o política.

Otros investigadores han ampliado los tipos de uso analizados (Cárdenas ...[et al], 2022). Por ejemplo, Behn (2002) propuso ocho formas de utilizar la información de evaluación en el proceso de diseño de políticas: para valoración, como mecanismo de control, para presupuestación, como elemento de motivación, de celebración, de aprendizaje, de mejoramiento y de promoción. De otra parte, Patton (1988) resalta la importancia del aprendizaje de los participantes cuando se involucran en las evaluaciones. Askim (2007) sugirió cinco tipos de uso de la información proveniente de las evaluaciones: oportunista, disfrazado, desconocido, de limpieza y circunstancial. Otros propósitos adicionales de la información de las evaluaciones se pueden observar en Moynihan y Landuyt (2009), Van Dooren ...[et al] (2010), Kroll (2012) y Sandison (2006).

Finalmente, se considera importante señalar que este documento asume que los gerentes públicos toman decisiones utilizando una racionalidad intencional, pero sus capacidades cognitivas están acotadas y la información con la que cuentan es incompleta (March, 1994). Es decir, cuando los individuos abordan problemas públicos,

los límites de información, conocimiento, tiempo y recursos, lo hacen con una racionalidad limitada (Simon, 1945), lo cual afecta el proceso de toma de decisiones (Lindblom, 1959).

1.2 Factores que inciden en el uso de la información evaluativa

Aunque existen múltiples agrupaciones sobre los factores que influyen en el uso de la información de las evaluaciones, este análisis considera las cuatro siguientes:

- **Tecnocráticos/racionales.** Según este enfoque el problema del uso limitado de las evaluaciones se presenta porque no se realiza una adecuada recolección y análisis de la información y, por tanto, no se puede obtener un insumo propicio para la toma de decisiones (Taylor, 2011b). En este sentido, los aspectos que pueden influir en el uso de la información evaluativa se refieren a la veracidad, la oportunidad y la confiabilidad (credibilidad) de la información producida (Landry ...[et al], 2003). En la medida en que la información disponible sea confiable, veraz y utilice metodologías robustas, se incrementará la posibilidad de utilizarla en la toma de decisiones (Poister y Streib, 1999; Taylor, 2011a). Shaxson y Boaz (2020) consideran que la calidad de la información incide en forma positiva en el uso de la información. Cousins ...[et al] (2014) llegan a la conclusión de que la oportunidad, la credibilidad de los hallazgos y la forma de comunicarlos son factores fundamentales para la que la información de las evaluaciones sea usada por los gerentes públicos.

- **Individuales.** Según estos factores, la percepción del funcionario público sobre la utilidad de la información de las evaluaciones, así como su orientación hacia el servicio público, incide en el uso de dichos documentos. Para Behn (2002) una de las barreras psicológicas que impiden un mayor uso de las evaluaciones en la toma de decisiones es el miedo que estas provocan en los funcionarios. Además, ha sido documentada la forma en que las personas con un mayor compromiso hacia los resultados de las organizaciones encuentran satisfacción en la búsqueda de mecanismos para mejorar los desempeños de estas (Perry y Hondeghe, 2008).

- **Organizacionales.** Estos factores hacen referencia a los incentivos en la organización para usar información de las evaluaciones. Incluye el uso por parte de los jefes y/o compañeros y las normas al interior de la organización que estimulen su uso. Al respecto, Rutter

(2012) considera que la baja demanda de la información evaluativa se debe a una falta de cultura que promueva este tipo de prácticas al interior de la organización. Organizaciones que se caracterizan por tener ambientes donde hay creencias y normas compartidas, alto empoderamiento, mayor participación y aprendizaje, tienden a incrementar el uso de la información de seguimiento y evaluación (Moynihan y Hawes, 2010).

- Políticos. Estos factores se refieren al apoyo de altos funcionarios del gobierno, el congreso y la ciudadanía del uso de la información. Para De Lancer Julnes y Holzer (2001) el éxito sostenible en el uso de la información de desempeño depende del apoyo continuado y sistemático de los funcionarios electos y del público en general. Los grupos de interés (*stakeholders*) internos y externos pueden jugar un papel fundamental y tener una influencia positiva en el momento de adoptar y usar información evaluativa en la toma de decisiones.

La interacción de todos estos factores podría incidir en el grado de utilización de la evidencia en la toma de decisiones (Cárdenas ... [et al], 2022), debido a que los procesos de implementación de las políticas públicas son particularmente susceptibles a las percepciones de los gerentes públicos, como lo sugieren académicos como Pressman y Wildavsky (1984), Weiss (1979), Lipsky (1980), O'Toole Jr. (2004) y Tummers y Bekkers (2014).

En este documento se analiza la relación de algunos factores racionales (métodos cuantitativos, cualitativos y mixtos) y el tipo de uso de dicha información, tomando como punto de partida la hipótesis sugerida por Leviton y Hughes (1981) que indica que los gerentes prefieren información cualitativa para la toma de decisiones, más que la cuantitativa que los evaluadores puedan brindarles.

2. Metodología

Para conocer el uso de las evaluaciones, se construye una base de datos con información sobre un conjunto de evaluaciones. Las fuentes consultadas son la Dirección de Seguimiento y Evaluación de Políticas Públicas del Departamento Nacional de Planeación, entidad encargada de coordinar el Sistema Nacional de Evaluación de Gestión y Resultados (Sinergia) en Colombia (Ley 152 de 1994), y las bibliotecas de diferentes universidades y algunas organizaciones no gubernamentales. En el caso de las bibliotecas

universitarias se consideraron las evaluaciones correspondientes a tesis de pregrado y posgrado de las universidades de los Andes, Externado de Colombia y Nacional, principalmente. Si bien es cierto que algunas de estas evaluaciones no se diseñaron con el fin de servir de insumo en la toma de decisiones por parte de los administradores públicos, este análisis pretende revisar tipos de uso diferentes al tradicional, es decir, diferentes al uso instrumental.

La búsqueda se limitó al período 2008-2019 y a los sectores de educación, salud, inclusión social, agricultura, transporte e infraestructura, y vivienda a nivel nacional. En total se encontró información para 251 evaluaciones. Una de las razones para escoger estos sectores es tener una visión enfocada más en los temas sociales (educación, salud e inclusión social) y en los aspectos económicos (agricultura, vivienda y transporte e infraestructura). Es importante resaltar que inicialmente la base de datos tenía un total de registros superior, pero se filtraron aquellos documentos que se definían como “línea de base” al considerarlos más un insumo para una evaluación posterior de un programa o políticas que una evaluación en estricto sentido.

La unidad de análisis de este estudio son las evaluaciones, más que los documentos donde se encontraban plasmadas dichas evaluaciones. Esto resulta relevante, porque los documentos que incluían más de una evaluación (ej. una evaluación de resultados y una de impacto) se contabilizaron en la base de datos de forma separada.

En la elaboración de la base de datos se tuvieron en cuenta aspectos como el nombre, el año de publicación, el sector, el autor institucional, el autor individual, el tipo de contratación, el tipo de evaluación según el tiempo o su objeto, los actores relevantes, el implementador de los resultados propuestos, la metodología empleada, el tipo de uso, los usuarios de la evaluación y el tipo de recomendaciones, entre otras cuestiones.

Se clasificaron dentro de la categoría “uso instrumental” los documentos evaluativos que esperaban que la información se usara para tomar decisiones y hacer cambios directos e inmediatos en el diseño del programa o la política, en su implementación o en la asignación de recursos. Ahora bien, aquellas evaluaciones que promovían una utilización de los análisis para procesos de rendición de cuentas, para hacer diagnósticos y comprender mejor cómo funciona un programa

Si bien el análisis del uso potencial o esperado de las evaluaciones no hace referencia al uso real de dichos estudios, sí permite describir y mostrar la intencionalidad de quienes diseñaron las evaluaciones y puede tener una alta correlación con los tipos de uso de la información.

o política pública, o para desarrollar planes de acción a futuro, se clasificaron en la categoría de “uso conceptual”.

Para realizar las regresiones con Mínimos Cuadrados Ordinarios (MCO) se consideraron como variables dependientes los tipos de uso de la información instrumental, conceptual y ambos tipos de uso. En esta última variable se incluyen las evaluaciones que registraron de forma explícita su propósito de utilizar la información tanto de forma instrumental como conceptual. Como variables independientes se escogieron los métodos utilizados en las evaluaciones: cuantitativos, cualitativos y mixtos. Se emplearon las siguientes variables de control: i) el tipo de contratación de las evaluaciones; ii) el tipo de evaluación según su objeto (insumos, procesos, resultados e impactos); iii) el tipo de evaluación según su tiempo (*ex ante*, concomitante y *ex post*); y iv) los sectores administrativos donde se implementaron las evaluaciones.

Por otra parte, aunque el uso simbólico se considera importante en esta investigación, no forma parte del análisis, porque no aparece de manera explícita como propósito en ninguna evaluación el empleo de los resultados para efectuar alguna acción relacionada con ese uso. Generalmente, el uso simbólico se estudia empleando métodos cualitativos y de forma *ex post* a la evaluación.

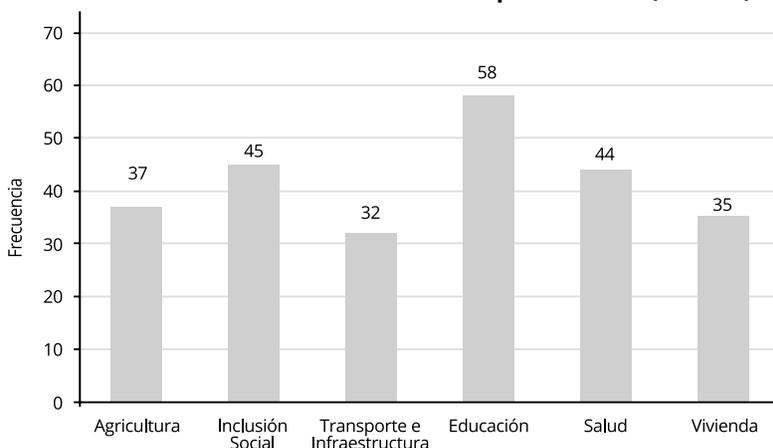
Cabe aclarar que en este ejercicio académico se emplea el uso “potencial”, porque en la compilación de la base de datos no se logra obtener mayor información sobre el uso “efectivo” de las evaluaciones. Si bien el análisis del uso potencial o esperado de las evaluaciones no hace referencia al uso real de dichos estudios, sí permite describir y mostrar la intencionalidad de quienes diseñaron las evaluaciones y puede tener una alta correlación con los tipos de uso de la información estudiados. Así, la información del uso potencial se obtuvo a partir de los propósitos definidos de forma *ex ante* en las evaluaciones que se estudiaron.

3. Resultados

3.1 Hallazgos descriptivos

De las 251 evaluaciones registradas en la base de datos 23,1% corresponde al sector educación (58); 17,9% es parte del sector inclusión social (45) y 17,5% del sector salud (44). El sector con menor número de evaluaciones en el estudio fue transporte e infraestructura (12,7%). Esto sugiere que la mayor cantidad de evaluaciones en el país corresponde a los sectores sociales (Gráfico 1).

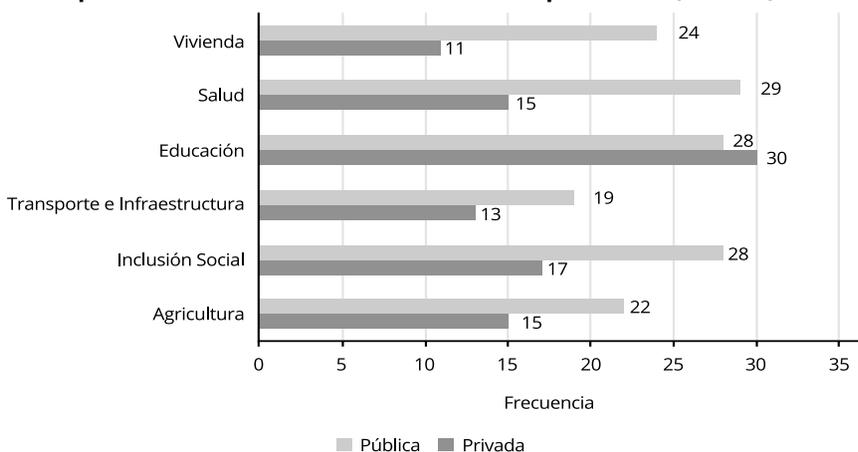
Gráfico 1
Documentos de evaluación analizados por sectores (N = 251)



Fuente: elaboración propia.

Las evaluaciones suelen ser contratadas por el sector público (59,8%) en mayor medida que por el sector privado (40,2%), siendo educación la excepción a la regla, pues se observa una mayor cantidad de evaluaciones contratadas por el sector privado (Gráfico 2). Cabe resaltar que en el sector privado no se habla propiamente de contratos en todos los casos, pues una cantidad considerable de evaluaciones corresponde a tesis de estudiantes universitarios.

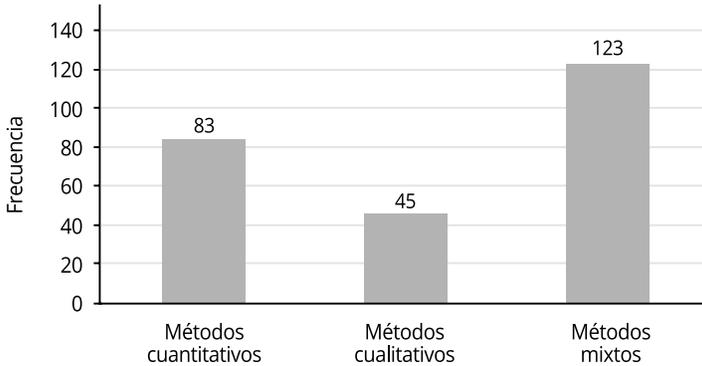
Gráfico 2
Tipo de contratación de las evaluaciones por sector (N = 251)



Fuente: elaboración propia.

Se observa que la metodología de métodos mixtos es utilizada en la mayoría de las evaluaciones (49%), seguida por los métodos cuantitativos (49%), seguida por los métodos cuantitativos y, en último lugar, por los métodos cualitativos (Gráfico 3).

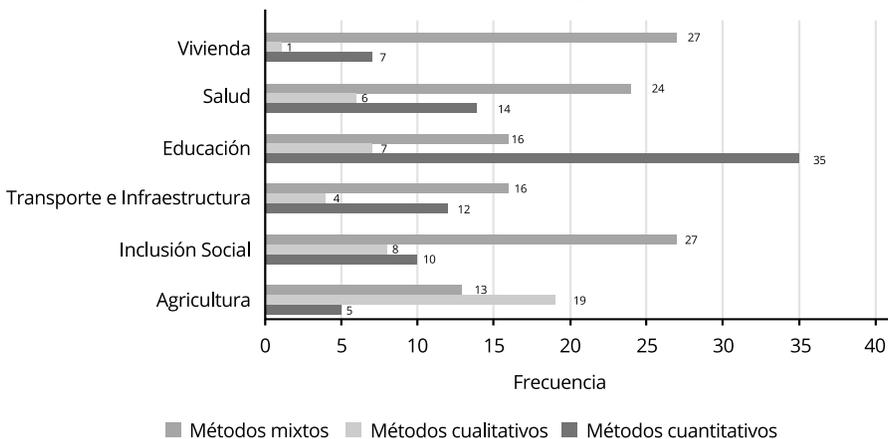
Gráfico 3
Metodología de las evaluaciones de política (N = 251)



Fuente: elaboración propia.

Agricultura es el sector con mayor número de evaluaciones empleando métodos cualitativos (19), mientras que educación tiene un mayor número de evaluaciones que utilizan métodos cuantitativos (35). Los sectores de vivienda e inclusión social tienen un mayor número de evaluaciones con métodos mixtos (27). Ver Gráfico 4.

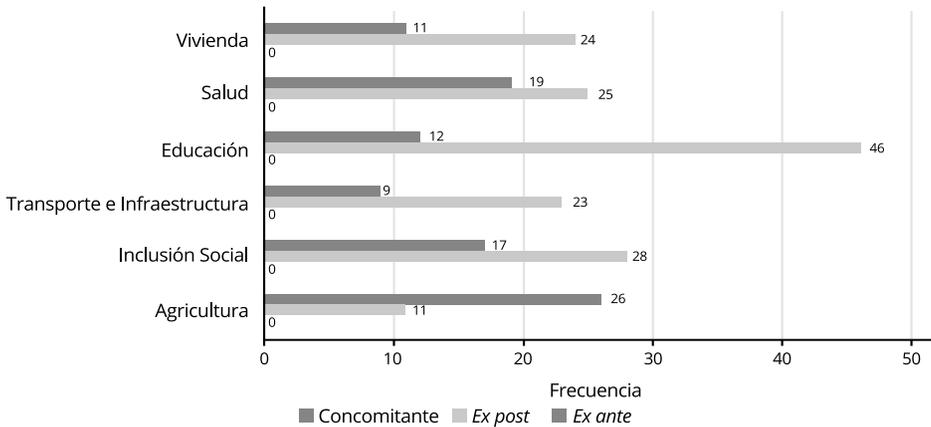
Gráfico 4
Sector discriminado por metodología (N = 251)



Fuente: elaboración propia.

Según su temporalidad (Ordóñez-Matamoros ...[et al], 2018), en todos los sectores la mayoría de las evaluaciones son *ex post* (62,5%), excepto en agricultura cuyas evaluaciones son concomitantes (Gráfico 5). En la compilación de información no se encontraron evaluaciones *ex ante*.

Gráfico 5
Tipo de evaluación (temporalidad) discriminado por sector (N=251)



Fuente: elaboración propia.

Las evaluaciones estudiadas suelen ser contratadas en mayor medida por el sector público, la metodología de métodos mixtos es la más utilizada y la mayoría en todos los sectores son evaluaciones *ex post*.

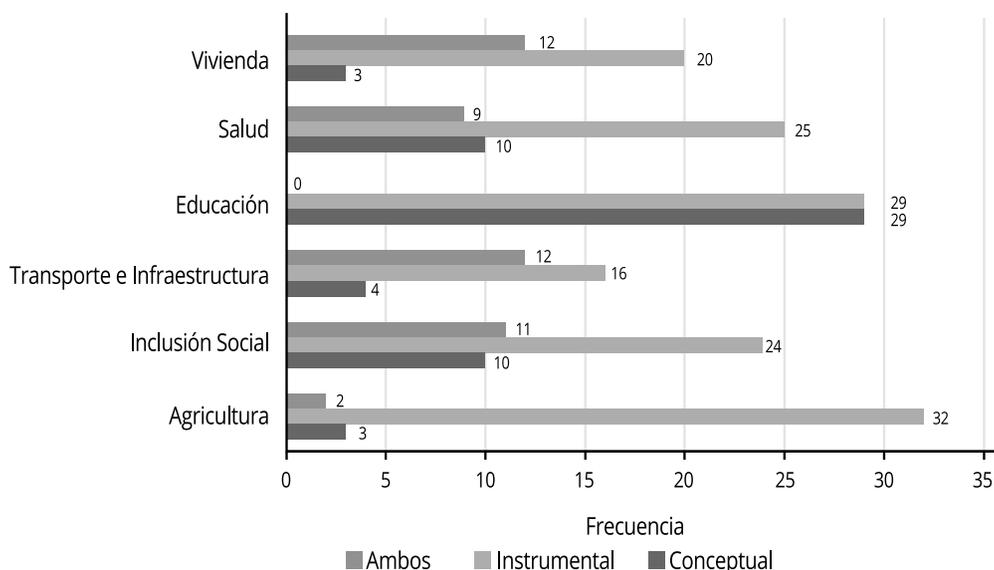
En resumen, se puede señalar que las evaluaciones estudiadas suelen ser contratadas en mayor medida por el sector público, la metodología de métodos mixtos es la más utilizada y la mayoría en todos los sectores son evaluaciones *ex post*. Por otra parte, agricultura es el sector con mayor número de evaluaciones con métodos cualitativos, mientras que educación tiene un mayor número de evaluaciones que utilizan métodos cuantitativos.

3.2 Resultados específicos sobre el tipo de uso potencial de las evaluaciones

Se hace referencia al uso potencial, porque en la compilación de la base no se pudo obtener más información sobre el uso efectivo de las evaluaciones. La información del uso potencial se obtuvo a partir de los propósitos que se definieron de forma *ex ante* en las mismas evaluaciones. Los tipos de uso de la información evaluativa que se tuvieron en cuenta para clasificar las evaluaciones fueron conceptual e instrumental. El uso simbólico no se tuvo en cuenta, porque no está explícito en ningún documento evaluativo.

La mayoría de las evaluaciones tiene un uso potencial instrumental (58,2%), en la medida en que formula recomendaciones específicas sobre cómo mejorar el programa, mientras que 23,5% presenta un uso potencial de tipo conceptual. Algunas evaluaciones, además de formular recomendaciones para modificar la política, son elaboradas para rendir cuentas. Por eso, se clasifican aparte, pues tienen elementos de los dos tipos de uso (18,3%). El tipo de uso potencial instrumental de las evaluaciones es bastante alto en todos los sectores. Se destacan los sectores de agricultura y educación en este tipo de uso. Por su parte, el tipo de uso conceptual también es bastante alto en el sector educación (Gráfico 6).

Gráfico 6
Sector de las políticas discriminado por tipo de uso (N = 251)

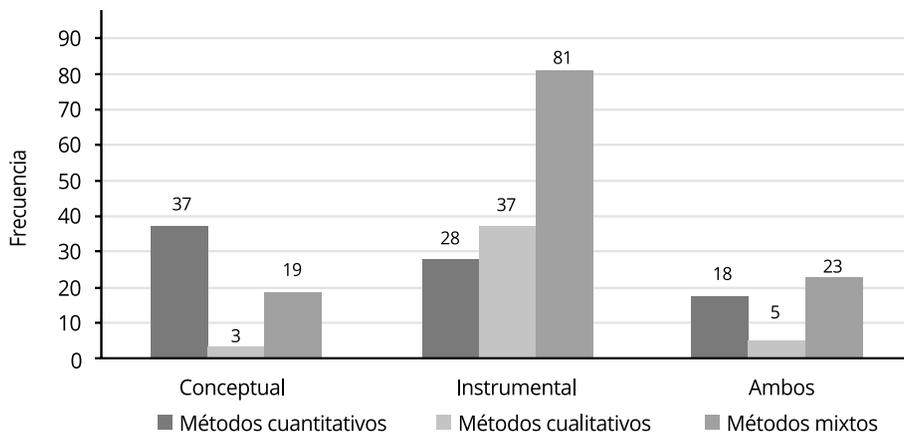


Fuente: elaboración propia.

Teniendo en cuenta la metodología empleada, la mayoría de las evaluaciones que utiliza métodos mixtos y cualitativos tiene un potencial uso instrumental. En cambio, gran parte de las evaluaciones realizadas con métodos cuantitativos tiene un potencial uso conceptual (Gráfico 7).

Gráfico 7

Tipo de uso de las evaluaciones discriminado por metodología (N = 251)

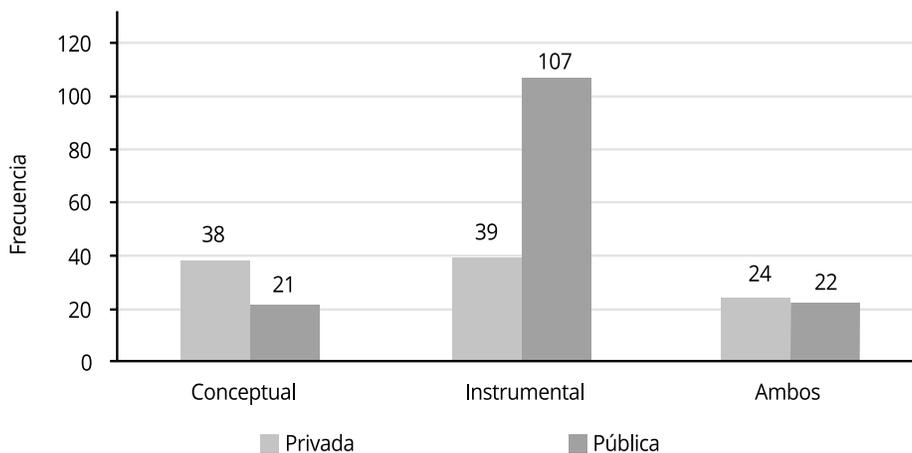


Fuente: elaboración propia.

Una gran mayoría de evaluaciones provenientes del sector público tiene un potencial uso instrumental, lo cual tiene sentido si se considera que los tomadores de decisiones contratarían dichas evaluaciones buscando saber si deben continuar o expandir el programa o no y si es posible mejorar su efectividad (Gráfico 8). El sector privado, a su vez, no muestra un interés marcado por algún tipo de uso.

Gráfico 8

Tipo de uso de las evaluaciones discriminado por contratación (N = 251)

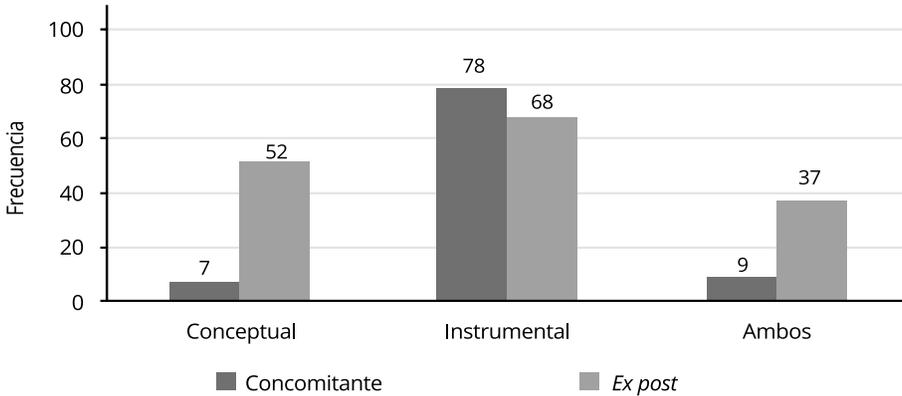


Fuente: elaboración propia.

Se observa también que, independientemente de la temporalidad, las evaluaciones tienden a tener un potencial uso instrumental más que conceptual. No obstante, se destaca que las evaluaciones concomitantes tienden a preferir, en gran medida, un uso potencial instrumental (Gráfico 9).

Gráfico 9

Tipo de uso discriminado por tipo de evaluación (temporalidad) (N = 251)

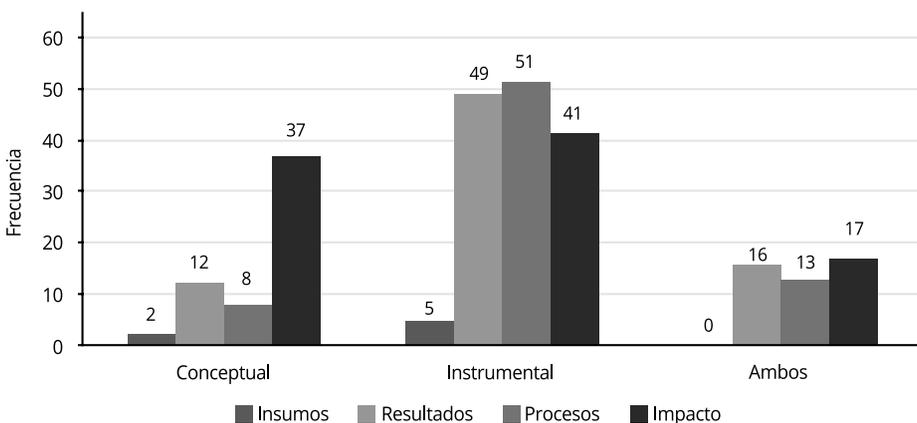


Fuente: elaboración propia.

Como se señala previamente, se observa que el tipo de uso potencial tiende a ser instrumental y que el tipo de evaluación según el objeto no influye mucho. Sin embargo, cuando se pretende efectuar un uso conceptual, se prefieren las evaluaciones de impacto (Gráfico 10).

Gráfico 10

Tipo de uso de discriminado por tipo de evaluación (objeto) (N = 251)



Fuente: elaboración propia.

En Colombia se diseñan evaluaciones con el objetivo de tomar decisiones directas e inmediatas sobre los programas y las políticas, es decir, para un uso instrumental, especialmente en los sectores de agricultura y educación.

En resumen, en Colombia se diseñan evaluaciones con el objetivo de tomar decisiones directas e inmediatas sobre los programas y las políticas, es decir, para un uso instrumental, especialmente en los sectores de agricultura y educación. En segundo lugar se elaboran evaluaciones para temas relacionados con rendición de cuentas y para entender mejor cómo funcionan los programas y las políticas públicas, es decir, para un uso conceptual, siendo también bastante alto en el sector educación. Adicionalmente, la mayoría de las evaluaciones contratadas por el sector público que emplearon métodos mixtos y cualitativos, tienen un potencial uso instrumental, mientras que gran parte de las evaluaciones realizadas con métodos cuantitativos tiene un potencial uso conceptual.

3.3 Modelo de regresión sobre los usos potenciales de las evaluaciones

Como se menciona en la sección sobre la metodología, para desarrollar las regresiones con MCO se tuvo en cuenta como variable dependiente los tipos de uso de la información instrumental, conceptual y ambos tipos de uso. Como variables independientes se escogieron los métodos utilizados en las evaluaciones: cuantitativos, cualitativos y mixtos. Y como variables de control se emplearon el tipo de contratación de la evaluación, el tipo de evaluación según su objeto, el tipo de evaluación según su tiempo, y los sectores administrativos donde se implementaron dichas evaluaciones.

Dada la naturaleza de las variables con las que se trabaja (binarias), es necesario partir de unas categorías base para poder correr una regresión. La evaluación "base" sería una evaluación privada, realizada con métodos cuantitativos, de impacto, *ex post* y de educación. Cualquier cambio en alguna de esas características se puede interpretar como un aumento o una disminución en la probabilidad de que la evaluación sea usada de forma instrumental o conceptual. Esto se expresa de la siguiente manera:

(1) *instrumental*_{*i*}

$$= \beta_0 + \beta_1 \text{ método_cuali}_i + \beta_2 \text{ método_mixto}_i + \beta_3 \text{ pública}_i + \beta_4 \text{ insumos}_i + \beta_5 \text{ proceso}_i + \beta_6 \text{ resultados}_i + \beta_7 \text{ concomitante}_i + \beta_8 \text{ agricultura}_i + \beta_9 \text{ salud}_i + \beta_{10} \text{ transporte}_i + \beta_{11} \text{ inclusión}_i + \beta_{12} \text{ vivienda}_i + \varepsilon_i$$

(2) *conceptual*_{*i*}

$$= \beta_0 + \beta_1 \text{ método_cuali}_i + \beta_2 \text{ método_mixto}_i + \beta_3 \text{ pública}_i + \beta_4 \text{ insumos}_i + \beta_5 \text{ proceso}_i + \beta_6 \text{ resultados}_i + \beta_7 \text{ concomitante}_i + \beta_8 \text{ agricultura}_i + \beta_9 \text{ salud}_i + \beta_{10} \text{ transporte}_i + \beta_{11} \text{ inclusión}_i + \beta_{12} \text{ vivienda}_i + \varepsilon_i$$

(3) *ambos_i*

$$= \beta_0 + \beta_1 \text{ método_cuali}_i + \beta_2 \text{ método_mixto}_i + \beta_3 \text{ pública}_i + \beta_4 \text{ insumos}_i + \beta_5 \text{ proceso}_i + \beta_6 \text{ resultados}_i + \beta_7 \text{ concomitante}_i + \beta_8 \text{ agricultura}_i + \beta_9 \text{ salud}_i + \beta_{10} \text{ transporte}_i + \beta_{11} \text{ inclusión}_i + \beta_{12} \text{ vivienda}_i + \varepsilon_i$$

Cuando se analiza el uso potencial instrumental utilizando MCO se encuentra que hay una relación estadísticamente significativa entre esta variable dependiente y los métodos utilizados en la evaluación (Tabla 1). Es decir, cuando la evaluación emplea métodos cualitativos, se incrementa la probabilidad de uso instrumental en 28,6 puntos porcentuales (pp) con respecto a las evaluaciones que utilizan métodos cuantitativos. La relación entre el uso instrumental y las evaluaciones que emplean métodos mixtos es positiva, pero no es estadísticamente significativa.

Cuando la evaluación es pública, se incrementa la probabilidad de uso instrumental en 19 pp con respecto a las evaluaciones contratadas por el sector privado. Finalmente, si la evaluación es concomitante, se incrementa la probabilidad de uso instrumental en 26,1 pp con respecto a las evaluaciones *ex post*.

De acuerdo con la misma tabla, se observa una relación negativa y estadísticamente significativa entre los métodos cualitativos y el tipo de uso esperado conceptual. Si la evaluación emplea métodos cualitativos, se reduce la probabilidad de uso conceptual en 25,1 pp con respecto a las evaluaciones que desarrollan métodos cuantitativos. Asimismo, las evaluaciones del sector agricultura, transporte y vivienda presentan una menor probabilidad de uso conceptual con respecto al sector educación en 26,7; 32,9 y 37,1 pp, respectivamente.

Por último, cuando se analizan las relaciones estadísticamente significativas entre ambos tipos de uso (conceptual e instrumental) y algunas variables de control, se observa que si las evaluaciones se efectúan en los sectores de salud, transporte, inclusión y vivienda, se registra un incremento de probabilidad de ambos usos (instrumental y conceptual) en relación con el sector educación en 23,9; 38,5; 27,9 y 37,4 pp, respectivamente.

Tabla 1
Estimación de los coeficientes de los modelos (1), (2) y (3) usando MCO

Variables	(1) Uso instrumental	(2) Uso conceptual	(3) Ambos usos
Métodos cualitativos	0.286*** (0.103)	-0.251*** (0.0887)	-0.0344 (0.0847)

Tabla 1 (continuación)
Estimación de los coeficientes de los modelos (1), (2) y (3) usando MCO

Variables	(1) Uso instrumental	(2) Uso conceptual	(3) Ambos usos
Métodos mixtos	0.185** (0.07829)	-0.131* (0.0672)	-0.0541 (0.0641)
Contratación pública	0.190*** (0.0679)	-0.116** (0.0583)	-0.0739 (0.0556)
Evaluación de insumos	-0.168 (0.189)	0.262 (0.163)	-0.0939 (0.155)
Evaluación de procesos	-0.131 (0.101)	0.0692 (0.0864)	0.0619 (0.0824)
Evaluación de resultados	-0.0610 (0.0816)	0.0195 (0.0700)	0.0415 (0.0668)
Evaluación concomitante	0.261*** (0.0842)	-0.135* (0.0723)	-0.126* (0.0690)
Sector agricultura	0.139 (0.102)	-0.267*** (0.0875)	0.128 (0.0835)
Sector salud	-0.0401 (0.0918)	-0.199** (0.0788)	0.239*** (0.0752)
Sector transporte	-0.0552 (0.101)	-0.329*** (0.0868)	0.385*** (0.0828)
Sector inclusión	-0.0795 (0.0917)	-0.200** (0.0787)	0.279*** (0.0751)
Sector vivienda	-0.00250 (0.104)	-0.371*** (0.0894)	0.374*** (0.0853)
Constante	0.298*** (0.0665)	0.635*** (0.0571)	0.0675 (0.0545)
Observaciones	251	251	251
R-cuadrado	0.243	0.245	0.174

Errores estándar en paréntesis.

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Fuente: elaboración propia.

No obstante, cuando se emplea MCO se pueden generar algunos inconvenientes. Uno es que las probabilidades predichas para cada unidad de observación no están acotadas entre 0 y 1. Por tal razón, los valores por fuera de ese intervalo no tienen mucho sentido cuando se trata de probabilidades como es el caso del uso instrumental, donde las probabilidades predichas se encuentran entre 0.09 y 1.11.

El uso instrumental (de las evaluaciones) tiene una relación estadísticamente significativa y positiva con los métodos cualitativos, la contratación pública y la evaluación concomitante.

Otro obstáculo es que la varianza del error en el modelo de regresión condicional sobre las variables independientes no sea constante, es decir, que haya heterocedasticidad. Para probar la presencia de este problema se estiman los modelos (1), (2) y (3) con MCO, se obtienen los residuales cuadrados para cada observación y se ejecuta regresión de los residuales cuadrados contra las variables independientes de los modelos (Wooldridge, 2009). Los estadísticos F (92.89) y los p-valores correspondientes (0.00) indican que hay heterocedasticidad. Para corregirla se emplean errores estándar robustos al ejecutar las regresiones de cada uno de los modelos. Esta modificación se muestra en la Tabla 2, donde se observa que la magnitud y el signo de los coeficientes no se modifican, pero sí su significancia estadística.

El uso instrumental tiene una relación estadísticamente significativa y positiva con los métodos cualitativos, la contratación pública y la evaluación concomitante (Tabla 2). Por su parte, el uso conceptual tuvo una relación estadísticamente significativa y negativa con los métodos cualitativos, el sector agricultura, transporte y vivienda. Así, se observa una relación estadísticamente significativa y positiva entre el uso “doble” y los sectores de salud, transporte, inclusión y vivienda, respectivamente.

Tabla 2
Estimación de los coeficientes de los modelos (1), (2) y (3)
usando MCO y errores estándar robustos

Variables	(1) Uso instrumental	(2) Uso conceptual	(3) Ambos usos
Métodos cualitativos	0.286*** (0.103)	-0.251*** (0.0886)	-0.0344 (0.0880)
Métodos mixtos	0.185** (0.0849)	-0.131* (0.0772)	-0.0541 (0.0728)
Contratación pública	0.190*** (0.0707)	-0.116* (0.0619)	-0.0739 (0.0624)
Evaluación de insumos	-0.168 (0.114)	0.262* (0.158)	-0.0939 (0.108)
Evaluación de procesos	-0.131 (0.0976)	0.0692 (0.0813)	0.0619 (0.0973)

Tabla 2 (continuación)
Estimación de los coeficientes de los modelos (1), (2) y (3)
usando MCO y errores estándar robustos

Variables	(1) Uso instrumental	(2) Uso conceptual	(3) Ambos usos
Evaluación de resultados	-0.0610 (0.0870)	0.0195 (0.0759)	0.0415 (0.0783)
Evaluación concomitante	0.261*** (0.0832)	-0.135* (0.0601)	-0.126 (0.0775)
Sector agricultura	0.139 (0.0932)	-0.267*** (0.0871)	0.128** (0.0537)
Sector salud	-0.0401 (0.0973)	-0.199** (0.0977)	0.239*** (0.0667)
Sector transporte	-0.0552 (0.101)	-0.329*** (0.0907)	0.385*** (0.0892)
Sector inclusión	-0.0795 (0.103)	-0.200** (0.0906)	0.279*** (0.0738)
Sector vivienda	-0.00250 (0.116)	-0.371*** (0.0873)	0.374*** (0.0905)
Constante	0.298*** (0.0742)	0.635*** (0.0724)	0.0675** (0.0289)
Observaciones	251	251	251
R-cuadrado	0.243	0.245	0.174

Errores estándar robustos en paréntesis.

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Fuente: elaboración propia.

En este caso resulta útil emplear modelos *probit* o *logit* para calcular los efectos marginales. Usando la misma especificación de los modelos (1), (2) y (3) se hacen regresiones y se calculan los efectos marginales del modelo *probit*, como se aprecia en las Tablas 3, 4 y 5.

Tabla 3
Regresión (Uso Instrumental) *Probit*

Variabes	dy/dx	Std. Err.	z	$P > z$	I	95%	C.I.
Métodos cualitativos	0.291	0.090	3.240	0.001	0.115	0.467	0.179
Métodos mixtos	0.179	0.088	2.040	0.042	0.007	0.352	0.490

Tabla 3 (continuación)
Regresión (Uso Instrumental) Probit

Variables	<i>dy/dx</i>	<i>Std. Err.</i>	<i>z</i>	<i>P > z</i>	<i>[-</i>	<i>95%</i>	<i>C.I.</i>
Contratación pública	0.220	0.079	2.770	0.006	0.064	0.376	0.598
Evaluación de insumos	-0.183	0.289	-0.630	0.527	-0.748	0.383	0.028
Evaluación de procesos	-0.172	0.127	-1.360	0.174	-0.420	0.076	0.287
Evaluación de resultados	-0.068	0.096	-0.710	0.478	-0.257	0.120	0.307
Evaluación concomitante	0.305	0.089	3.440	0.001	0.131	0.479	0.375
Sector agricultura	0.218	0.107	2.040	0.041	0.009	0.428	0.147
Sector salud	-0.047	0.109	-0.430	0.667	-0.261	0.167	0.175
Sector transporte	-0.062	0.122	-0.510	0.612	-0.301	0.177	0.127
Sector inclusión	-0.095	0.109	-0.870	0.387	-0.309	0.120	0.179
Sector vivienda	0.011	0.120	0.090	0.929	-0.225	0.246	0.139

Fuente: elaboración propia.

Como se establece previamente, los métodos cualitativos, la contratación pública y la evaluación concomitante presentan una relación estadísticamente significativa (al 5%) y positiva con el uso potencial instrumental (Tabla 3). En este sentido, cuando la evaluación emplea métodos cualitativos se incrementa la probabilidad de uso instrumental en 29,1 pp con respecto a las evaluaciones que utilizan métodos cuantitativos.

No obstante, a diferencia de las regresiones obtenidas con MCO, la relación entre el uso instrumental y las evaluaciones que emplean métodos mixtos es positiva, pero es estadísticamente significativa. Cuando la evaluación emplea métodos mixtos se incrementa la probabilidad de uso instrumental en 17,9 pp con respecto a las evaluaciones que utilizan métodos cuantitativos. A diferencia de las regresiones obtenidas con MCO, las evaluaciones en el sector agricultura también tienen una relación estadísticamente significativa y positiva con el uso instrumental de la información de las evaluaciones.

Tabla 4
Regresión (Uso Conceptual) *Probit*

Variabes	<i>dy/dx</i>	<i>Std. Err.</i>	<i>z</i>	<i>P > z</i>	<i>l</i>	95%	<i>C.I.</i>
Métodos cualitativos	-0.171	0.053	-3.260	0.001	-0.274	-0.068	0.179
Métodos mixtos	-0.099	0.064	-1.550	0.122	-0.224	0.027	0.490
Contratación pública	-0.115	0.063	-1.830	0.068	-0.239	0.008	0.598
Evaluación de insumos	0.419	0.265	1580	0.115	-0.101	0.938	0.028
Evaluación de procesos	0.099	0.109	0.910	0.362	-0.114	0.313	0.287
Evaluación de resultados	0.013	0.071	0.190	0.853	-0.126	0.152	0.307
Evaluación concomitante	-0.168	0.071	-2.370	0.018	-0.307	-0.029	0.375
Sector agricultura	-0.180	0.044	-4.070	0.000	-0.266	-0.093	0.147
Sector salud	-0.112	0.051	-2.190	0.028	-0.213	-0.012	0.175
Sector transporte	-0.185	0.038	-4.880	0.000	-0.259	-0.111	0.127
Sector inclusión	-0.132	0.049	-2690	0.007	-0.229	-0.036	0.179
Sector vivienda	-0.210	0.038	-5.570	0.000	-0.284	-0.136	0.139

Fuente: elaboración propia.

De forma similar a lo establecido con los resultados obtenidos con MCO, utilizando el modelo *probit* se encuentra que los métodos cualitativos y los sectores de agricultura, transporte y vivienda presentan una relación estadísticamente significativa (al 5%) y negativa con el uso conceptual. En este sentido, si la evaluación emplea métodos cualitativos se reduce la probabilidad de uso conceptual en 17,1 pp con respecto a las evaluaciones que desarrollan métodos cuantitativos. Además, las evaluaciones del sector agricultura, transporte y vivienda presentan una menor probabilidad de uso conceptual con respecto al sector educación en 18,0; 18,5 y 21,0 pp, respectivamente (Tabla 4).

No obstante, al utilizar el modelo *probit* se observan relaciones estadísticamente significativas y negativas adicionales tanto con la evaluación concomitante como con los sectores salud e inclusión (Tabla 4).

Tabla 5
Regresión (Doble uso) *Probit*

Variables	dy/dx	Std. Err.	z	$P > z$	$[-$	95%	$]$	C.I.
Métodos cualitativos	-0.014	1096	-0.010	0.990	-2163	2134	0.168	
Métodos mixtos	-0.011	0.849	-0.010	0.989	-1675	1652	0.492	
Contratación pública	-0.026	1827	-0.010	0.989	-3607	3556	0.598	
Evaluación de procesos	0.015	1113	0.010	0.989	-2165	2196	0.295	
Evaluación de resultados	0.008	0.561	0.010	0.989	-1092	1107	0.316	
Evaluación concomitante	-0.030	2241	-0.010	0.989	-4423	4363	0.365	
Sector agricultura	0.957	9253	0.100	0.918	-17179	19094	0.139	
Sector salud	0.980	4547	0.220	0.829	-7932	9893	0.180	
Sector transporte	0.993	1650	0.600	0.547	-2242	4228	0.131	
Sector inclusión	0.987	3207	0.310	0.758	-5300	7273	0.180	
Sector vivienda	0.993	1662	0.600	0.550	-2264	4251	0.135	

Fuente: elaboración propia.

En contraste con los hallazgos obtenidos con MCO, se observan resultados muy diferentes al emplear el modelo *probit* (y *logit*) utilizando como variable dependiente el doble uso potencial (conceptual e instrumental). De esta forma, ninguna variable independiente tiene relación estadísticamente significativa con aquellas evaluaciones que presentan un uso potencial tanto instrumental como conceptual (Tabla 5). En ese sentido, es posible que la variable construida para identificar la situación en la que las evaluaciones pretenden alcanzar dos tipos de uso en forma simultánea esté ocasionando las diferencias tan marcadas.

En resumen, en Colombia se diseñan evaluaciones para alcanzar un uso de tipo instrumental, especialmente en los sectores de agricultura y educación. En segundo lugar, se elaboran evaluaciones para lograr un uso de tipo conceptual. Asimismo, se encontró que el uso instrumental tiene una relación estadísticamente significativa y positiva con los métodos cualitativos, la contratación pública y las

evaluaciones concomitantes. Por otra parte, el uso conceptual tuvo una relación estadísticamente significativa y negativa con los métodos cualitativos, en los sectores de agricultura, transporte y vivienda.

Conclusiones y discusión

Con base en la información compilada sobre 251 documentos de evaluación de políticas entre 2008 y 2019 en seis sectores de Colombia -educación, salud, inclusión, agricultura, transporte y vivienda, se encuentra que cerca del 60% de las evaluaciones analizadas fueron contratadas por el sector público, 49% utiliza métodos mixtos y, de acuerdo con su objeto, 37,8% corresponde a evaluaciones de impacto y 30,6% a evaluaciones de resultados.

Adicionalmente, de acuerdo con la información que brinda la base de datos, y respondiendo a la primera pregunta del estudio, se encuentra que la información de las evaluaciones sí se utiliza potencialmente de forma instrumental y conceptual en los sectores analizados. Se observa que las evaluaciones se diseñan principalmente con el objetivo de tomar decisiones directas e inmediatas sobre los programas y las políticas, es decir, para un potencial uso instrumental. En segundo lugar se diseña para temas de rendición de cuentas, un potencial uso conceptual. Se evidencia que las evaluaciones no se diseñan para un uso exclusivo relacionado con la toma de decisiones de manera directa sobre las políticas públicas (uso instrumental), sino también para otros propósitos.

Cuando se emplean Mínimos Cuadrados Ordinarios (MCO) y errores estándar robustos para analizar los factores que pueden incidir en los potenciales usos de los resultados de dichas evaluaciones, se encontró que el uso instrumental tiene una relación estadísticamente significativa y positiva con los métodos cualitativos, la contratación pública y las evaluaciones concomitantes. En otras palabras, cuando las evaluaciones utilizan métodos cualitativos, son contratadas por el sector público y son de tipo concomitante, es más probable que los resultados de dichos análisis se utilicen para tomar decisiones de forma directa e inmediata sobre los programas y las políticas públicas.

Aunque, como se ha descrito en el documento, hay poca literatura empírica al respecto, el hallazgo que relaciona los métodos cualitativos con el tipo de uso instrumental corrobora lo encontrado por Leviton y Hughes (1981) cuando afirman que los administradores prefieren información cualitativa (para la toma de decisiones) a la

Los tomadores de decisiones prefieren análisis más descriptivos, que narren historias, que se enfoquen en las interpretaciones, las experiencias y su significado.

Un uso para realizar cambios relacionados con la operación o implementación de los programas públicos (en vez de un uso para hacer cambios más profundos relacionados con la asignación de presupuesto y/o la reformulación de políticas públicas, es más frecuente, porque no suscita mayor debate político o social.

cuantitativa que los evaluadores brindan. En este sentido, se puede sugerir que los tomadores de decisiones prefieren análisis más descriptivos, que narren historias, que se enfoquen en las interpretaciones, las experiencias y su significado.

En este sentido, de acuerdo con March (1994), en un mundo de uso de información de política pública que se encuentra enmarcado por factores que fomentan la ambigüedad, el desarrollo de significado es mucho más importante que los resultados. Según Rutter (2012), los políticos prefieren guiar sus decisiones de política pública por aspectos netamente subjetivos (construcción de significado) más que por información "objetiva".

De esta forma, el significado no se establece para la toma de decisiones, sino que la toma de decisiones se realiza para desarrollar significado (March, 1994). El desarrollo de significado es la actividad organizacional central en la toma de decisiones. Por tal motivo, si se desea que la información de las evaluaciones permee la toma de decisiones, se requiere construir significado (March, 1994) y desarrollar mecanismos adecuados para esta clase de escenarios de alta ambigüedad a través de la argumentación, la persuasión (Majone, 2014) y el juego del lenguaje (Wittgenstein, 2014). Cualquier intento por mejorar la toma de decisiones debe concebir a la decisión como un instrumento de significado.

Por otra parte, el hecho de encontrar una asociación positiva entre las evaluaciones concomitantes (o que se enfocan en la implementación de una política o programa) y el uso potencial instrumental, en lugar de evaluaciones de tipo *ex post*, llama la atención. Al parecer, la toma de decisiones se pretende realizar sobre aspectos operativos o de procesos de la política o programa, más que sobre decisiones profundas relacionadas, por ejemplo, con la asignación de presupuesto o con la expansión o eliminación de dicho programa o política. En este sentido, parece necesario distinguir entre un uso para realizar cambios relacionados con la operación o implementación de los programas públicos, de un uso para hacer cambios más profundos relacionados con la asignación de presupuesto y/o la reformulación de políticas públicas. Como lo sugiere González (2016), el primer tipo de uso es más frecuente, porque se trata de aspectos muy concretos que no suscitan mayor debate político o social y/o no están inmersos en ambientes organizacionales con amplios problemas de ambigüedad.

No obstante, aunque el uso instrumental de la información de las evaluaciones se ha concentrado en cambios puntuales operacionales y de implementación de los programas públicos, no se puede desdeñar

Si se va a utilizar información de las evaluaciones para procesos de rendición de cuentas, los métodos cuantitativos pueden ofrecer información más puntual y concreta sobre los avances de un programa o política que los métodos cualitativos.

Las organizaciones prefieren utilizar la información solo para mostrar resultados cuantitativos favorables por el miedo de una cobertura negativa por parte de los medios de comunicación.

este logro ya que, como lo ha mostrado la literatura de implementación de política pública (Pressman y Wildavsky, 1984), pequeñas modificaciones en temas fundamentales pueden generar cambios sustanciales que permitan alcanzar los propósitos de algún programa (González, 2016). Asimismo, como lo plantea Lindblom (1959), en ausencia de un cálculo cierto de las consecuencias que producirán decisiones profundas, lo realmente racional para un gobierno es realizar correcciones limitadas y sucesivas a sus decisiones pasadas (incrementalismo).

Algunos autores consideran que más importante que los datos y los métodos empleados en las evaluaciones es tener un equipo de funcionarios capaz de analizar la información evaluativa de forma oportuna y veraz (Berman y Wang, 2002).

Por su parte, el uso conceptual tuvo una relación estadísticamente significativa y negativa con los métodos cualitativos, en los sectores de agricultura, transporte y vivienda. En otras palabras, es más probable que el uso conceptual se incremente cuando la evaluación emplea métodos cuantitativos y la evaluación se desarrolla en el sector de educación, por ejemplo, para efectuar procesos de rendición de cuentas o para entender mejor un programa o una política.

Una hipótesis sugerida para el primer resultado es que si se va a utilizar información de las evaluaciones para procesos de rendición de cuentas, los métodos cuantitativos pueden ofrecer información más puntual y concreta sobre los avances de un programa o política que los métodos cualitativos. En este sentido, Taylor (2011a) sugiere que las organizaciones usan más la información de desempeño para requerimientos externos (rendición de cuentas) que para mejoras internas. La misma investigadora sugiere que las organizaciones prefieren utilizar la información solo para mostrar resultados cuantitativos favorables por el miedo de una cobertura negativa por parte de los medios de comunicación (Taylor, 2011a). Adicionalmente, se puede sugerir que el sector de educación requiere mostrar resultados cuantitativos de forma sistemática y continua no solo al interior del gobierno, sino a otros actores fuera del mismo, incluyendo a la ciudadanía.

Finalmente, se observa una relación estadísticamente significativa y positiva entre el uso potencial "doble" y los sectores de salud, transporte, salud y vivienda, respectivamente. Es decir, el uso instrumental y conceptual de las evaluaciones se incrementa de forma simultánea en los sectores de salud, transporte, inclusión y vivienda con respecto al sector de educación.

Cuando se emplean modelos *probit* para analizar la información se obtienen resultados similares para los usos potenciales instrumental

y conceptual. No obstante, para el primer tipo de uso se encuentra, además, que los métodos mixtos y las evaluaciones en el sector agricultura también tienen una relación estadísticamente significativa y positiva con el uso instrumental de la información de las evaluaciones. Con respecto al uso conceptual se encontró que la evaluación concomitante y los sectores de salud e inclusión social también muestran una relación estadísticamente significativa y negativa.

Ahora bien, al emplear estos modelos *probit* se encuentra que ninguna variable independiente tiene relación estadísticamente significativa con las evaluaciones que presentan un uso potencial doble, tanto instrumental como conceptual. Una posibilidad para corregir este problema es dejar por fuera de la regresión las evaluaciones que tienen dos usos potenciales de forma simultánea, o introducirlas dos veces, una vez por cada uso registrado.

De otra parte, utilizar el concepto de uso potencial puede generar limitaciones en el análisis. Inicialmente, es probable que la mayoría de los entrevistados señalen que quieren diseñar evaluaciones para modificar políticas y programas y, en menor grado, para efectuar diagnósticos o para temas de rendición de cuentas. Este deseo al elaborar las evaluaciones se puede contrastar en gran medida cuando se observa el uso efectivo de las mismas. En este orden de ideas, cuando se pueda efectuar un análisis sobre el uso efectivo de dichas evaluaciones se podrá tener información más veraz sobre el verdadero uso instrumental, conceptual y simbólico de la información.

En este sentido, para tener un mayor entendimiento sobre el uso efectivo (no solo potencial) de la información de las evaluaciones en investigaciones posteriores, así como para avanzar en el conocimiento de otros tipos de uso, como el simbólico, y los factores que inciden en estos, se requiere profundizar en la investigación utilizando, por ejemplo, entrevistas semiestructuradas o encuestas a los usuarios directos de documentos evaluativos.

Finalmente, para avanzar en este campo de investigación se requiere profundizar más en términos conceptuales y en la operacionalización de los tipos de uso de la información proveniente de las evaluaciones (instrumental, conceptual y simbólico), así como en los factores que inciden en dichos usos. Por otra parte, con el fin de fortalecer el análisis comparativo del estudio, se considera relevante poner a prueba los resultados de este trabajo con análisis en otros países de Latinoamérica que han tenido importantes avances en materia de Gestión Pública Orientada a Resultados.

Bibliografía

- Aiello, E.; Donovan, C.; Duque, E.; Fabrizio, S.; Flecha, R.; Holm, P.; Molina, S.; Oliver, E.; y Reale, E. (2021), "Effective Strategies that Enhance the Social Impact of Social Sciences and Humanities Research", en *Evidence and Policy*, Vol. 17 N° 1, pp. 131-146, <https://doi.org/10.1332/174426420X15834126054137>.
- Alkin, M. y S. Taut (2002), "Unbundling Educational Use", en *Studies in Educational Evaluation*, Vol. 29 N° 1, pp. 1-12, [http://dx.doi.org/10.1016/S0191-491X\(03\)90001-0](http://dx.doi.org/10.1016/S0191-491X(03)90001-0).
- Appleton-Dyer, S.; Clinton, J.; Carswell, P.; y McNeill, R. (2012), "Understanding Evaluation Influence within Public Sector Partnerships: a Conceptual Model", en *The American Journal of Evaluation*, Vol. 33 N° 4, pp. 532-546, <https://doi.org/10.1177/1098214012447672>.
- Arellano, D. y Hernández, J. (2014), "De la Torre de Marfil a la pertinencia. ¿La investigación académica en administración pública es útil para la toma de decisiones gubernamental? Un estudio exploratorio del caso mexicano", México, Centro de Investigación y Docencia Económicas (Dossier Académico).
- Askim, J. (2007), "How Do Politicians Use Performance Information? An Analysis of the Norwegian Local Government Experience", en *International Review of Administrative Sciences*, Vol. 73 N° 3, pp. 453-472.
- Balthasar, A. (2009), "Institutional Design and Utilization of Evaluation. A Contribution to a Theory of Evaluation Influence Based on Swiss Experience", en *Evaluation Review*, Vol. 33 N° 3, pp. 226-256.
- Behn, R. (2002), "The Psychological Barriers to Performance Management: or Why Isn't Everyone Jumping on the Performance-Management Bandwagon?", en *Public Performance and Management Review*, Vol. 25 N° 1, pp. 5-25.
- Berman, E. y Wang, X. (2002), "Performance Measurement in U.S. Counties: Capacity for Reform", en *Public Administration Review*, Vol. 60 N° 5, pp. 409-420.
- Boyer, J. F. y Langbein, L. I. (1991), "Factors Influencing the Use of Health Evaluation Research in Congress", en *Evaluation Review*, Vol. 15 N° 5, pp. 507-532.
- Bridgeland, J. y Orszag, P. (2013), "Can Government Play Moneyball?", en *The Atlantic*, Boston, July-August, <https://www.theatlantic.com/magazine/archive/2013/07/can-government-playmoneyball/309389/>, 06-12-2018.
- Burr, E. (2009), "Evaluation Use and Influence among Project Directors of State GEAR UP Grants", Knoxville, The University of Tennessee, Ph.D. Dissertation, http://trace.tennessee.edu/utk_graddiss/15.

- Cairney, P. (2019), "The UK Government's Imaginative Use of Evidence to Make Policy", en *British Politics*, Vol. 14 N°1, pp. 1-22, <https://doi.org/10.1057/s41293-017-0068-2>.
- Cárdenas, S.; González, J.; y Ramírez, E. (2022), *Aprender de la evidencia. Usos de la información en la gestión de organizaciones educativas*, Bogotá, Universidad de los Andes; Centro de Investigación y Docencia Económicas.
- Cardozo, M. (2006), "La evaluación de políticas y programas públicos. El caso de los programas de desarrollo social en México", México, Cámara de Diputados; Porrúa.
- _____ (2013a), "Políticas públicas: los debates de su análisis y evaluación", en *Andamios*, Vol. 10 N° 21, pp. 39-59.
- _____ (2013b), "De la evaluación a la reformulación de políticas públicas", en *Política y Cultura*, N° 40, pp. 123-149, <http://www.redalyc.org/articulo.oa?id=26728947007>.
- Chelimsky, E. (2006), "The Purposes of Evaluation in a Democratic Society", en *The Sage Handbook of Evaluation*, I. Shaw, J. Greene y M. Mark (eds.), London, Sage Publishing.
- Choi, Y. y Woo, H. (2021), "Understanding Diverse Types of Performance Information Use: Evidence from an Institutional Isomorphism Perspective", en *Public Management Review*, Vol. 24 N° 12, pp. 1-20, <https://doi.org/10.1080/14719037.2021.1955953>.
- Cousins, J.; Goh, S.; Elliott, C.; y Bourgeois, I. (2014), "Framing the Capacity to Do and Use Evaluation", en *New Directions for Evaluation*, Vol. 2014 N° 141, pp. 7-23, <https://doi.org/10.1002/ev.20076>.
- Cunill, N. y Ospina, S. (2003), "Una agenda de investigación sobre la evaluación de los resultados de la gestión pública", en *Evaluación de resultados para un gestión pública moderna y democrática. Experiencias latinoamericanas*, N. Cunill y S. Ospina (eds.), Caracas, CLAD.
- De Lancer Julnes, P. (2006), "Performance Measurement an Effective Tool for Government Accountability? The Debate Goes on", en *Evaluation*, Vol. 12 N° 2, pp. 219-235.
- De Lancer Julnes, P. y Holzer, M. (2001), "Promoting the Utilization of Performance Measures in Public Organizations: an Empirical Study of Factors Affecting Adoption and Implementation", en *Public Administration Review*, Vol. 61 N° 6, pp. 693-708.
- Fetterman, D. (1996), "Empowerment Evaluation: an Introduction to Theory and Practice", en *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability*, D. M. Fetterman, S. J. Kaftarian, y A. Wandersman (eds.), Thousand Oaks, Sage.

- Fleischer, D. y Christie, C. (2009), "Evaluation Use: Results from a Survey of U.S. American Evaluation Association Members", en *American Journal of Evaluation*, Vol. 30 N° 2, pp. 158-175, <http://dx.doi.org/10.1177/1098214008331009>.
- French, R. (2018), "Lessons from the Evidence on Evidence-Based Policy", en *Canadian Public Administration*, Vol. 61 N° 3, pp. 425-442, <https://doi.org/10.1111/capa.12295>.
- González, J. (2016), "Por qué se usa la información de seguimiento y evaluación de políticas públicas", México, Centro de Investigación y Docencia Económicas. Tesis para optar por el título de Doctor en Políticas Públicas.
- Halachmi, A. (2002), "Performance Measurement, Accountability, and Improved Performance", en *Public Performance and Management Review*, Vol. 25 N° 4, pp. 370-374.
- Heinrich, C. y Marschke, G. (2010), "Incentives and their Dynamics in Public Sector Performance Management Systems", en *Journal of Policy Analysis and Management*, Vol. 29 N° 1, pp. 183-208.
- Henry, G. (2000), "Why Not Use?", en *New Directions for Evaluation*, Vol. 2000 N° 88 (Special Issue: The Expanding Scope of Evaluation Use), pp. 85-98.
- Hojlund, S. (2014), "Evaluation Use in the Organizational Context: Changing Focus to Improve Theory", en *Evaluation*, Vol. 20 N° 1, pp. 26-43.
- Hyryläinen, E. y Viinamäki, O. (2008), "The Implications of the Rationality of Decision-Makers on the Utilization of Evaluation Findings", en *International Journal of Public Administration*, Vol. 31 Nos. 10-11, pp. 1223-1240, <https://doi.org/10.1080/01900690801973311>.
- Johansson, T. y Siverbo, S. (2009), "Explaining the Utilization of Relative Performance Evaluation in Local Government: a Multi-Theoretical Study Using Data from Sweden", en *Financial Accountability and Management*, Vol. 25 N° 2, pp. 197-224.
- Kroll, A. (2012), "Why Public Managers Use Performance Information: Concepts, Theory, and Empirical Analysis", Potsdam, The University of Potsdam. Doctoral Thesis.
- Kumar, M.; Taegtmeier, M.; Madan, J.; Ndimba, S.; Chikaphupha, K.; Kea, A.; y Barasa, E. (2020), "How Do Decision-Makers Use Evidence Incommunity Health Policy and Financing Decisions? A Qualitative Study and Conceptual Framework in Four African Countries", en *Health Policy and Planning*, Vol. 35 N° 7, pp. 799-809.
- Lammert, J.; Heinemeier, S.; y Fiore, T. (2017), *Effectively Communicating Evaluation Findings*, Rockville, Westat.

- Landry, R.; Lamari, M.; y Amara, N. (2003), "The Extent and Determinants of the Utilization of University Research in Government Agencies", en *Public Administration Review*, Vol. 63 N° 2, pp. 192-205.
- Lapuente, V. y Van de Walle, S. (2020), "The Effects of New Public Management on the Quality of Public Services", en *Governance*, Vol. 33 N° 3, pp. 461-475, <https://doi.org/10.1111/gove.12502>.
- Ledermann, S. (2012), "Exploring the Necessary Conditions for Evaluation Use in Program Change", en *American Journal of Evaluation*, Vol. 33 N° 2, pp. 159-178, <https://doi.org/10.1177/1098214011411573>.
- Leviton, L. y Hughes, E. (1981), "Research on the Utilization of Evaluations: a Review and Synthesis", en *Evaluation Review*, Vol. 5 N° 4, pp. 525-548.
- Lindblom, C. (1959), "The Science of Muddling through", en *Public Administration Review*, Vol. 19 N° 2, pp. 79-88.
- Lipsky, M. (1980), *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*, New York, Russell Sage Foundation.
- Luna, C. (2016), "La evaluación de políticas públicas centrada en el uso de resultados y de aprendizajes", en *RevIISE: Revista de Ciencias Sociales y Humanas*, Vol. 8 N° 8, pp. 7-14.
- Majone, G. (2014), *Evidencia, argumentación y persuasión en la formulación de políticas*, México, Fondo de Cultura Económica.
- March, J. (1994), *A Primer on Decision Making: How Decisions Happen*, New York, The Free Press; Simon and Schuster.
- Mark, M. y Henry, G. (2004), "The Mechanisms and Outcomes of Evaluation Influence", en *Evaluation*, Vol. 10 N° 1, pp. 35-57.
- McDonnell, L. y Weatherford, M. S. (2020), *Evidence, Politics, and Education Policy*, Cambridge, Harvard Education Press.
- Mikkelsen, M. F.; Petersen N. B. G; y Bjørnholt, B. (2021), "Broadcasting Good News and Learning from Bad News: Experimental Evidence on Public Managers' Performance Information Use", en *Public Administration*, Vol. 1 N° 19, <https://doi.org/10.1111/padm.12762>.
- Moynihan, D. (2005), "Goal-Based Learning and the Future of Performance Management", en *Public Administration Review*, Vol. 65 N° 2, pp. 203-216.
- _____ (2009), "Through a Glass, Darkly: Understanding the Effects of Performance Regimes", en *Public Performance and Management Review*, Vol. 32 N° 4, pp. 592-603.
- Moynihan, D. y Hawes, D. (2010), "The Organizational Environment and Performance Information Use", paper prepared for the European Group of Public Administration Meeting in Toulouse, France, September 8-11.

- _____ (2012), "Responsiveness to Reform Values: the Influence of the Environment on Performance Information Use", en *Public Administration Review*, Vol. 72 N° 1, pp. 95-105.
- Moynihan, D. e Ingraham, P. (2001), "When Does Performance Information Contribute to Performance Information Use? Putting the Factors in Place; Working Paper", New York, Syracuse University. Campbell Public Affairs Institute.
- Moynihan, D. y Landuyt, N. (2009), "How Do Public Organizations Learn? Bridging Cultural and Structural Perspectives", en *Public Administration Review*, Vol. 69 N° 6, pp. 1097-1105.
- Moynihan, D. y Pandey, S. (2010), "The Big Question for Performance Management: Why Do Managers Use Performance Information?", en *Journal of Public Administration Research and Theory*, Vol. 20 N° 4, pp. 849-866.
- Newcomer, K. (1997), "Using Performance Measurement to Improve Public and Nonprofit Programs", en *New Directions for Evaluation*, N° 75, pp. 5-14.
- Newman, J.; Cherney, A.; y Head, B. (2017), "Policy Capacity and Evidence-Based Policy in the Public Service", en *Public Management Review*, Vol. 19 N° 2, pp. 157-174, <http://dx.doi.org/10.1080/14719037.2016.1148191>.
- OCDE (2002), *Glossary of Key Terms in Evaluation and Results-Based Management*, Paris, OCDE.
- _____ (2020), *Improving Governance with Policy Evaluation. Lessons from Country Experiences*, Paris, OCDE, <https://doi.org/10.1787/89b1577d-en>.
- Ongaro, M. y Andreotti, M. (2022), "Non-Empirical Uncertainties in Evidence-Based Decision-Making", en *Perspective on Science*, Vol. 30 N° 2, pp. 305-320.
- Ordóñez-Matamoros, G.; Centeno, J.; Santander Ramírez, D.; Llanos Congote, O.; Sierra Daza, M; y Remolina Pulido, J. (2018), "La evaluación ejecutiva y la toma de decisiones de política pública: el caso de la E2+SGR en Colombia", en *OPERA*, N° 22, pp. 53-79, <https://doi.org/10.18601/16578651.n22.04>.
- Oszlak, O. (2013), "Gobierno abierto: hacia un nuevo paradigma de gestión pública", s.l., Red de Gobierno Electrónico de América Latina y el Caribe (Colección de Documentos de Trabajo sobre e-Gobierno; N° 5).
- O'Toole Jr., L. (2004), "The Theory-Practice Issue in Policy Implementation Research", en *Public Administration*, Vol. 82 N° 2, pp. 309-329, <https://doi.org/10.1111/j.0033-3298.2004.00396.x>.
- Owen, J. y Lambert, F. (1998), "Evaluation and the Information Needs of Organizational Leaders", en *American Journal of Evaluation*, Vol. 19 N° 3, pp. 355-365.

- Pallett, H. (2020), "The New Evidence-Based Policy: Public Participation between 'Hard Evidence' and Democracy in Practice", en *Evidence and Policy*, Vol. 16 N° 2, pp. 209-227, <https://doi.org/10.1332/174426419X15704985880872>.
- Parkhurst, J. (2016), "*The Politics of Evidence. From Evidence-Based Policy to the Good Governance of Evidence*", London, Routledge, <https://doi.org/10.4324/9781315675008>.
- Patton, M. (1988), "The Evaluator's Responsibility for Utilization", en *Evaluation Practice*, Vol. 9 N° 2, pp. 5-24.
- Pawson, R.; Wong, G; y Owen, L. (2011), "Known Knowns, Known Unknowns, Unknown Unknowns: the Predicament of Evidence-Based Policy", en *American Journal of Evaluation*, Vol. 32 N° 4, pp. 518-546.
- Peck, L. y Gorzalski, L. (2009), "An Evaluation Use Framework and Empirical Assessment", en *Journal of Multidisciplinary Evaluation*, Vol. 6 N° 12, pp. 139-156.
- Perry, J. y Hondeghem, A. (2008), *Motivation in Public Management: the Call of Public Service*, Oxford, Oxford University Press.
- Poister, T. y Streib, G. (1999), "Performance Measurement in Municipal Government: Assessing the State of the Practice", en *Public Administration Review*, Vol. 59 N° 4, pp. 325-335.
- Preskill, H. (2004), "The Transformational Power of Evaluation: Passion, Purpose, and Practice", en *Evaluation Roots*, Marvin C. Alkin (ed.), California, Sage Publications, <https://dx.doi.org/10.4135/9781412984157.n23>.
- Pressman, J. y Wildavsky, A. (1984), *Implementation. How Great Expectations in Washington Are Dashed in Oakland*, Berkeley, University of California Press. 3. ed.
- Rickinson, M.; De Bruin, K.; Walsh, L.; y Hall, M. (2017), "What Can Evidence-Use in Practice Learn from Evidence-Use in Policy?", en *Educational Research*, Vol. 59 N° 2, pp. 173-189, <https://doi.org/10.1080/00131881.2017.1304306>.
- Rossi, P.; Lipsey, M.; y Freeman, H. (2004), *Evaluation: a Systematic Approach*, Thousand Oaks, Sage Publications. 7. ed.
- Rutter, J. (2012), *Evidence and Evaluation in Policy Making: a Problem of Supply or Demand?*, London, Institute for Government, <http://www.instituteforgovernment.org.uk/publications/evidence-and-evaluation-policy-making>.
- Saliterer, I. y Korac, S. (2013), "Performance Information Use by Politicians and Public Managers for Internal Control and External Accountability Purposes", en *Critical Perspectives on Accounting*, Vol. 24 Nos. 7-8, pp. 502-517.

- Sandison, P. (2006), "The Utilisation of Evaluations", en *ALNAP Review of Humanitarian Action: Evaluation Utilisation*, J. Mitchel (ed.), London, Active Learning Network for Accountability and Performance in Humanitarian Action, pp. 89-144.
- Shaxson, L. (2019), "Uncovering the Practices of Evidence-Informed Policy-Making", en *Public Money and Management*, Vol. 39 N° 1, pp. 46-55, <https://doi.org/10.1080/09540962.2019.1537705>.
- Shaxson, L. y Boaz, A. (2020), "Understanding Policymakers' Perspectives on Evidence Use as a Mechanism for Improving Research-Policy Relationships", en *Environmental Education Research*, Vol. 27 N° 4, pp. 518-524, <https://doi.org/10.1080/13504622.2020.1764505>.
- Shulha, L. y Cousins, J. (1997), "Evaluation Use: Theory, Research, and Practice Since 1986", en *American Journal of Evaluation*, Vol. 18 N° 3, pp. 195-208.
- Simon, H. (1945), *The Administrative Behavior*, New York, Free Press.
- Spekle, R. y Verbeeten, F. (2009), "The Use of Performance Measurement Systems in the Public Sector: Effects on Performance", paper presented at the AAA 2009 Management Accounting Section (MAS) Meeting, <https://ssrn.com/abstract=1162242>.
- Stufflebeam, D. (1983), "The CIPP Model for Program Evaluation", en *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, G. F. Madaus, M. S. Scriven y D. L. Stufflebeam (eds.), Boston, Kluwer-Nijhoff, pp. 117-141.
- Taylor, J. (2011a), "Factors Influencing the Use of Performance Information for Decision Making in Australian State Agencies", en *Public Administration*, Vol. 89 N° 4, pp. 1316-1334, <https://doi.org/10.1111/j.1467-9299.2011.02008.x>.
- _____ (2011b), "Strengthening the Link between Performance Measurement and Decision Making", en *Public Administration*, Vol. 89 N° 3, pp. 860-878.
- Tummers, L. y Bekkers, V. (2014), "Policy Implementation, Street-Level Bureaucracy, and the Importance of Discretion", en *Public Management Review*, Vol. 16 N° 4, pp. 527-547.
- UNEG (2016), "Normas y estándares de evaluación", Nueva York, Naciones Unidas. Grupo de Evaluación.
- Van Dooren, W.; Bouckaert, G.; y Halligan, J. (2010), *Performance Management in the Public Sector*, London, Routledge.
- Van de Walle, S. y Van Dooren, W. (2009), "How Is Information Used to Improve Performance in the Public Sector? Exploring the Dynamics of Performance Information", en *Connecting Knowledge and Performance in Public Services: from Knowing to Doing*, K. Walshe, G. Harvey y P. Jas (eds.), Oxford, Oxford University Press, <http://dx.doi.org/10.2139/ssrn.1491644>.

- Weiss, C. H. (1979), "The Many Meanings of Research Utilization", en *Public Administration Review*, Vol. 39 N° 5, pp. 426-431.
- _____ (1982), "Policy Research in the Context of Diffuse Decision Making", en *The Journal of Higher Education*, Vol. 53 N° 6, pp. 619-639.
- Wholey, J. S.; Hatry, H. P.; y Newcomer, K. E. (2010), *Handbook of Practical Program Evaluation*, San Francisco, Jossey-Bass.
- Wittgenstein, L. (2014), *Recherches Philosophiques*, Paris, Gallimard.
- Wooldridge, J. M. (2009), "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables, en *Economics Letters*, Vol. 104 N° 3, pp. 112-114.
- Zaltsman, A. (2010), *The Multiple Uses of Performance Information throughout the Budgetary Process: Insights from Chile's Performance-Based Budgeting System*, s.l., s.e.