
Cuadernos del CEDEOP

CENTRO DE ESTUDIOS DEL ESTADO

Y

LAS ORGANIZACIONES PÚBLICAS

N.º 28

Año 2 | 26 de noviembre de 2023



Dexter Fernandelli: 2023, Inspiração dos seis quadrantes sobre Gropius

IADCOM
Instituto de Investigaciones en Administración,
Contabilidad y Métodos Cuantitativos para la Gestión

• 1913 • 2023 • 110 Aniversario
.UBAeconómicas
FACULTAD DE CIENCIAS ECONÓMICAS

UNIVERSIDAD DE BUENOS AIRES

FACULTAD DE CIENCIAS ECONÓMICAS

AUTORIDADES DE LA FACULTAD VINCULADAS

CON LOS INSTITUTOS DE INVESTIGACIÓN

Decano

Dr. Ricardo Pahlen Acuña

Secretario de Investigación

Mg. Adrián Ramos

Directora del Instituto de Investigaciones

en Administración, Contabilidad y Métodos Cuantitativos
para la Gestión – IADCOM

Dra. María Teresa Casparri

Director del Centro de Estudios sobre el Estado
y las Organizaciones Públicas

Alejandro M. Estévez, Ph.D.



Staff

Director de la publicación:

Alejandro M. Estévez, Ph.D.

Comité Editorial:

Cristian Pliscoff Varas - Universidad Católica de Chile (Chile)

Cristina Díaz -Universidad Nacional de Entre Ríos (Argentina)

Denis Proulx - Université du Québec (Canadá)

Dora Bonardo –Universidad Nacional del Comahue (Argentina)

Eva Sørensen -Universidad Roskilde (Dinamarca)

Evert Vedung - Universidad de Uppsala (Suecia)

Jale Tosun – Universidad de Heidelberg (Alemania)

J. Ignacio Criado - Universidad Autónoma de Madrid (España)

Joan Subirats - Universidad Autónoma de Barcelona (España)

Jorge Etkin - Universidad de Buenos Aires (Argentina)

Leonardo Schvarstein - Universidad de Buenos Aires (Argentina)

Norberto H. Góngora - Universidad Nacional de La Plata (Argentina)

Omar Guerrero - Universidad Nacional Autónoma de México (México)

Patricia Domench – Universidad Nacional de Avellaneda (Argentina)

Diseño gráfico: Edwin H. McDonald (UBA)

ISSN: 2796-8200 // Publicación arbitrada confeccionada por el CEDEOP. // Corresponde exclusivamente a los autores la responsabilidad por los conceptos expuestos en los artículos firmados, de lo cual debe inferirse que el CEDEOP puede compartir las opiniones vertidas o no. // Se autoriza la reproducción de los textos incluidos en la revista, con la necesaria mención de la fuente.

Centro de Estudios del Estado y las Organizaciones Públicas (CEDEOP);
Facultad de Ciencias Económicas – Universidad de Buenos Aires.
Avenida Córdoba 2122 (C1120AAQ),
Ciudad Autónoma de Buenos Aires,
República Argentina
cedeop@fce.uba.ar
www.economicas.uba.ar

Una tipología de los doce usos de la evaluación

Evert Vedung¹

Resumen²

Utilizar los conocimientos de las evaluaciones es un mantra para los evaluadores. La solución habitual consiste en discernir los usos instrumental, conceptual y legitimador como usos de los productos de evaluación y considerar el uso de los procesos de evaluación como un cuarto tipo esencialmente distinto de los tres anteriores. Este artículo rompe con esta tipología de uso de 3+1 de dos maneras fundamentales. Una divergencia consiste en añadir los usos táctico, ritual y constitutivo de los productos a la tríada predominante de usos instrumental, conceptual y legitimador. Aún más fundamental es la segunda divergencia, en la que los seis usos se asocian no sólo al uso del producto, sino también al uso del proceso. Esto significa que los seis tipos de uso se presentan tanto como usos del producto como usos del proceso, lo que da lugar a un patrón de 6x2 en lugar del 3+1 de la tipología de usos de la evaluación. El resultado es una *tipología de los usos de la evaluación en doce partes*. Tras explicar la distinción sobre el uso del producto/uso del proceso de la evaluación, este *paper* analiza en detalle la tipología de usos de la evaluación en doce partes y brinda ejemplos prácticos de todos ellos para clarificar la cuestión.

Palabras clave: tipología de los usos de la evaluación en doce partes, modelo de ingeniería del uso de la evaluación, uso constitutivo, uso del producto, uso del proceso.

Abstract

Bringing evaluation insights to use is a mantra among evaluators. Their standard solution is to discern instrumental, conceptual, and legitimizing uses as utilizations of evaluation products and to regard use of evaluation processes as a fourth type essentially different from the previous three. In two fundamental ways, this article is breaking loose from this three-times-one typology of use. One deviation is adding tactical, ritual, and constitutive product uses to the prevailing triad of instrumental, conceptual, and legitimizing uses. Even more fundamental is the second divergence, where all six uses are joined not only to product use but to process use as well. This means that all six types of use will turn up in the form of both product use and process use, leading to a six-times-two pattern instead of the three-times-one-uses-of-evaluation typology. The result is the *Twelve-part typology of evaluation uses*. After elaborating on the evaluation product-use/evaluation process-use distinction, the paper deliberates on the twelve-part evaluation uses typology in detail and provides practical examples of them all for illumination.

Keywords: Twelve-part Typology of Evaluation Uses, Engineering Model of Evaluation Use, Constitutive Use, Product Use, Process Use

Fuente: Vedung, Evert, 2021, "The Twelve-Part Typology of Evaluation Uses", *Zeitschrift für Evaluation*, 20(1), 101-130. <https://doi.org./10.31244/zfe.2021.01.05>

Título en alemán: Die zwölfteilige Typologie der Evaluationsnutzungen

¹ Ph.D. in Political Science; Uppsala University, Sweden. Emeritus professor and director at Institute for Housing and Urban Research, Department of Government in Uppsala University.

² Este artículo fue traducido por la Mg. Guadalupe Laborde (UBA/USAL).

Quien, en la búsqueda de la ciencia, busca la utilidad práctica inmediata, puede estar seguro, en general, de que buscará en vano.

Hermann Von Helmholtz (1821–1894)

(Obtenido de: <https://gutezitate.com/autor/-hermann-von-helmholtz> [28.01.2020])

El uso de la evaluación suele referirse a uno de estos cuatro tipos de uso: instrumental, conceptual, simbólico y de proceso. El uso instrumental se refiere a los resultados de la evaluación para la toma de decisiones, como la continuación o el cambio del programa... El uso conceptual alude a la iluminación o aprendizaje general que se genera en la evaluación. El uso de la evaluación para justificar las opiniones existentes sin la intención real de obtener un beneficio del proceso o de los resultados se denomina uso simbólico. El uso del proceso se refiere a la instrucción o cambio motivado por el aprendizaje que se produce durante el proceso de una evaluación, más que por sus resultados.

Sarah Appleton-Dyer, Janet Clinton, Peter Carswell y Rob McNeill (2011)

(Obtenido de: <https://slideplayer.com/slide/9859342/> [01.10.2020])

Introducción

"Nuestra valía [como evaluadores] debe ser juzgada por la utilidad. Las evaluaciones son para ayudar a mejorar programas, para hacerlos más eficaces y proporcionar información para la toma de decisiones. Esa es la utilidad. Esa es nuestra responsabilidad".³

Michael Quinn Patton (1988:228) hizo esta afirmación sobre su utilidad, notoriamente formulada, durante su famosa disputa con Carol H. Weiss (1988) sobre la utilización, durante la conferencia de la *American Evaluation Society* de 1987, en Boston, MA, EE.UU. La utilidad es esencial, y lograr una utilidad es nuestro deber como evaluadores en ejercicio.⁴

El énfasis de Patton sobre su utilidad es compartido por numerosos evaluadores. Según Wolfgang Beywl, "la utilidad de una evaluación se manifiesta cuando los conocimientos, la información y las conclusiones son utilizados e influyen en la práctica sobre las acciones de las audiencias" (citado de Stockmann/Meyer 2013). Desde el comienzo de la evaluación de políticas públicas en la década de 1960, se espera que sus resultados se utilicen directa e inmediatamente para que las intervenciones evaluadas se vayan perfeccionando. Esto se denomina uso instrumental.

Pero desde el principio se alzaron voces frustradas lamentando el hecho de que los resultados de muchas evaluaciones excelentes se ignoraran o se utilizaran vergonzosamente y fueran malinterpretados o deliberadamente sesgados. Si los resultados de las evaluaciones se utilizan erróneamente o no se utilizan en absoluto, ¿qué sentido tiene dedicar tiempo y dinero a las evaluaciones? En vista de ello, teóricos de la evaluación como Weiss argumentaron que los críticos exigían demasiado; y que el uso conceptual o el uso en el pensamiento, pero no en la acción, era suficiente como fundamento y justificación para la realización de evaluaciones. Este punto de vista molestó y decepcionó a muchos, incluido Patton. También se descubrió que los resultados de la evaluación sólo se utilizaban para legitimar posiciones ya adoptadas. Con consternación, el uso legitimador fue acusado como un uso incorrecto de la evaluación. A lo largo de los años, los evaluadores han seguido debatiendo sobre su uso. La utilización de los procesos y resultados de la evaluación habían sido objeto de más deliberaciones que el uso equivalente de la ciencia social aplicada.

³Esta es una versión desarrollada, revisada y ampliada de mi capítulo en el Festschrift zum 60 de 2015. *Geburtstag von Reinhard Stockmann* (véase Vedung 2015).

⁴Para más información, véase Alkin/King (2006: 574).

1. Objetivo y estructura del artículo

El propósito general de este artículo sobre la utilidad de la evaluación es lógico y conceptual. En mi revisión de la bibliografía sobre el uso de la evaluación se abordan cuatro cuestiones (Tabla 1, preguntas 2-5). Antes de referirse a este cuarteto de preguntas, la lógica exige una definición fidedigna de lo que es la evaluación, por lo que empiezo preguntando: "Evaluación, ¿qué es eso?". (Tabla 1, pregunta 1). La sexta y última pregunta, "¿Por qué otro artículo más sobre la utilización de la evaluación?", pretende ser no sólo un resumen, sino también una exposición prospectiva de la cuestión del uso. Los objetivos específicos y la organización del siguiente texto están plasmados en el esquema de seis preguntas de la Tabla 1.

Tabla 1: Esquema de seis preguntas sobre la evaluación

1. Evaluación, ¿qué es eso?
2. ¿Qué aspectos de una evaluación se presupone que serán de utilidad?
3. ¿Qué tipos genéricos de uso de la evaluación existen?
4. ¿Sería fructífero contar con una combinación lógica de tipos genéricos y aspectos de la evaluación?
5. ¿Son todos los tipos genéricos de uso, aceptables de un modo ético, para los evaluadores serios?
6. ¿Por qué otro artículo más sobre la utilización de la evaluación?

La respuesta a la pregunta 1 consistirá en una definición minimalista del término evaluación. La respuesta a la pregunta 2 arrojará luz sobre dos aspectos de las evaluaciones que pueden utilizarse: los procesos y los productos. La adopción y utilización se desencadenan tanto por las actividades de evaluación (procesos) como por los resultados de las actividades de evaluación (productos). Aunque no es del todo original, es de esperar que mi línea divisoria, un tanto especial, entre estas dos propiedades resulte clarificadora para algunos evaluadores y teóricos.

Para responder a la pregunta 3 habrá que describir seis tipos genéricos de uso de la evaluación, incorporando el uso táctico, ritual y, fuera de lo común, el constitutivo, al conjunto de usos instrumentales, conceptuales y legitimadores comúnmente reconocidos. Precisaré y perfeccionaré cada uno de estos seis tipos genéricos para lógicamente mantenerlos separados y convertirlos en *herramientas* incisivas y fructíferas para pensar, argumentar, realizar investigaciones empíricas y seguir un razonamiento normativo.

La respuesta a la pregunta 4 es un rotundo sí. Una combinación lógica de los seis tipos genéricos de uso y de los dos aspectos de la evaluación sería fructífera y daría lugar a doce nuevos tipos de uso (Figura 1). También demostraré que tienen equivalentes en el mundo real. Por lo que sé, la bibliografía no contiene ninguna combinación de doce puntos de esta naturaleza, ilustrados con ejemplos empíricos (Vedung 2015; Dahler-Larsen 2012). Aquí tenemos lo que llamaré *la tipología de los usos de la evaluación en doce partes*.

¿Son aceptables de un modo ético para los evaluadores serios los seis tipos de uso de la evaluación y sus doce combinaciones? ¿O hay usos incorrectos de la evaluación claramente evidentes que los evaluadores deberían dejar de lado? Al introducir una

perspectiva temporal, sobre los usos de la evaluación en la pregunta 5 espero proyectar al menos, algo de luz sobre esta vieja cuestión.

Al responder a la pregunta 6, "¿Por qué otro artículo más sobre la utilización de la evaluación?", se pretende ofrecer algo más que un resumen convencional, proporcionando una visión concisa de porqué mi contribución es especial y con perspectivas de futuro.

2. Evaluación, ¿qué es eso?

La evaluación es el *proceso* de determinar el mérito de una entidad y el *producto* de esa determinación. Esta respuesta abstracta y general a la primera de las seis preguntas de la Tabla 1 capta el significado universal básico en inglés del término evaluación, como una importante actividad milenaria y el resultado de esa actividad en todos los esfuerzos humanos (cf. Scriven 2013: 178; también Alkin/King 2016:589).

En el contexto actual, la evaluación como la determinación del mérito se restringirá drásticamente al escrutinio y juicio de una familia especial de entidades: las intervenciones humanas para cambiar o impedir el cambio. Este significado restringido se denomina comúnmente evaluación de programas o simplemente evaluación, por conveniencia.

Mientras que algunos académicos se preocupan poco por definir formalmente los términos clave, yo prefiero hacerlo con expresiones esenciales. En mi opinión, el conocimiento conceptual es fundamental en las ciencias sociales. Además, siguiendo los pasos del gran politólogo italiano Giovanni Sartori (2009a: 89-91, 84-85; 2009b: 139-140), estoy a favor de las llamadas *definiciones mínimas*. Como consecuencia, enuncio las propiedades absolutamente necesarias que debe tener una entidad para ser calificada como evaluadora. Dado que la evaluación es un imán semántico (cf. Vedung 1997:3 y 1999:1), es importante aclarar el significado de la noción central en un contexto relacionado con la terminología académica. Para resumir una historia larga y complicada, propongo la siguiente definición mínima:

Evaluación: El proceso de valoración cuidadosa del mérito del contenido, la institucionalización, la implementación, el resultado y los efectos de la intervención, así como el producto resultante de este proceso, donde tanto el proceso como el producto tienen como fin ser utilizados en futuras situaciones de acción práctica.

La evaluación es una *valoración* (apreciación, juicio) interesada en determinar el mérito, es decir, el valor, de las intervenciones. El término se refiere a la valoración en dos sentidos: el proceso de valoración y el producto de esa valoración. Una persona puede menear un informe escrito en el aire y exclamar: "¡Esto es una evaluación!", mientras que simultáneamente confirma que: "¡La evaluación para este informe requirió 12 meses de duro trabajo!".

La evaluación es una valoración *minuciosa* que se ejecuta sistemática y meticulosamente sin ser necesariamente científica o académica.

Los objetivos a evaluar, o programas, se circunscriben a *intervenciones* de todas las clases, ya sean públicas o privadas. Intervenir es interponerse. Más específicamente, significa involucrarse en una situación como para alterar u obstaculizar una acción o desarrollo.

En una evaluación, las intervenciones pueden juzgarse por sus resultados en la sociedad y en la naturaleza. Para muchos teóricos de la evaluación, esto y sólo esto es la evaluación. Otros, entre los que me incluyo, no están de acuerdo y abogan por una visión más inclusiva, en la que la evaluación también incluya valoraciones limitadas al contenido

de la intervención y su institucionalización, además de cómo ésta puede influir en diversos niveles de organizaciones burocráticas y procesos de implementación, hasta las respuestas de los grupos objetivo. Es más, el término intervención incluye el mandato de generar algo, por ejemplo, nuevas propuestas de políticas. En ese caso, la implementación se refiere a los procesos creativos, a menudo intrincados, que implica la configuración de estas nuevas propuestas.

Al final y como mínimo, para poder recibir el calificativo de evaluación, los estudios que cumplen con todos los criterios anteriormente mencionados, también *deben servir para ser utilizados*, al menos en el hecho de ser notificados e informados por alguien, ya sean funcionarios, actores sociales u otros interesados en el (tipo de) intervención evaluada. Sin embargo, exigir como criterio definitorio la utilización real sería demasiado riguroso. Hay muchas evaluaciones buenas y oportunas que no se utilizan en absoluto, pero aún cumplen con los criterios para ser llamadas evaluación si tuvieran como fin ser utilizadas. Para reiterar, los estudios que cumplen todos los criterios enumerados, incluido el objetivo de ser utilizados, independientemente de si se abandonan por completo o no se les presta atención alguna, siguen siendo evaluaciones, aunque no sean utilizadas.

Para finalizar mi discurso sobre la definición de evaluación, metaevaluación, en sus dos sentidos: 1) análisis sistemático y resumido de los resultados y recomendaciones de varias evaluaciones y 2) auditorías evaluativas de procesos de evaluación puntuales y del funcionamiento de los sistemas de evaluación, ya que también es una evaluación, porque las evaluaciones en sí mismas son a su vez intervenciones, aunque sean de segundo orden.

Una última advertencia al lector antes de pasar a la pregunta 2 de la tabla 1 es que, aunque admitiendo en la definición que la evaluación puede ser dirigida a todo tipo de intervenciones, privadas, de la sociedad civil y públicas, el resto de este artículo se limitará a ejemplos prácticos de intervenciones públicas y de su implementación.

3. Uso del proceso y uso del producto

¿Qué aspectos de una evaluación se presupone que serán de utilidad? Los evaluadores experimentados probablemente responderían: productos y procesos, es decir, el uso de los resultados, recomendaciones y otros productos, y el uso de conocimientos que surgen de las actividades del proceso.

Cuando escribí la primera versión de este artículo en 2013-2014, para su inclusión en el *Festschrift* en honor a mi estimado colega Reinhard Stockmann (Vedung 2015), la opinión más común en la bibliografía era tratar los usos instrumental, conceptual y legitimador como usos fundamentales del producto y considerar el uso del proceso como un cuarto tipo de uso esencialmente diferente de los tres anteriores (véase el epígrafe de Appleton-Dyer et al. al principio de este artículo). He decidido romper con este patrón 3+1 y lo pondré en práctica en este artículo. Permítanme comenzar distinguiendo entre el uso del proceso y el uso del producto.

La división de uso del proceso/uso del producto se origina en la doble definición de la evaluación como el proceso de determinar el mérito de una intervención y el producto de ese proceso. Sin embargo, la distinción que voy a hacer aquí es más delicada de lo que algunos lectores puedan pensar en un principio. En primer lugar, el *uso del producto* abarca la utilización de los resultados y recomendaciones para la acción basadas en el informe final, pero que abarca no sólo éste, sino también la información y recomendaciones escritas, orales y visuales que complementan el informe final, por ejemplo, el contenido de las presentaciones en *Power-Point*, las explicaciones y ampliaciones orales complementarias, los seminarios, los

comunicados de prensa y las selecciones del informe final transmitidas a audiencias específicas. Además, y de forma más infrecuente, la categoría de producto también incluye la utilización de resultados y recomendaciones extraídas de los borradores del informe final y de los resultados provisionales escritos, orales o comunicados electrónicamente durante los procesos de evaluación, es decir, algunas veces, mucho antes de que se complete el informe final y se concluyan las actividades de evaluación (Chouinard et al., 2020).

El *uso del proceso*, por otra parte, abarca la utilización del aprendizaje derivado de participar en la evaluación propiamente dicha y de formar parte de sus actividades. El uso del proceso incluye cambios cognitivos, actitudinales y de comportamiento en los individuos, que resultan de su participación en los procedimientos de evaluación que, a su vez, directa o indirectamente, pueden generar reflexiones y discusiones que pueden producir cambios en la intervención evaluada o en alguna intervención similar. El uso del proceso consiste en aprender a pensar de forma evaluativa gracias a la participación en un proceso de evaluación y en aplicar ese aprendizaje en las deliberaciones y acciones sobre la intervención evaluada. Por ejemplo, ocurre cuando las personas que participan en una evaluación, más tarde admiten que el impacto de su propia intervención no se debió únicamente a los resultados, sino también a "haber atravesado el proceso de reflexión que la evaluación requería" (Patton 2012: 140.).

Sin embargo, el uso del proceso es más inclusivo. También abarca los conocimientos adquiridos de completar cuestionarios, ser entrevistado por los evaluadores, trabajar como miembro de un equipo, ayudar a determinar y realizar ajustes en las preguntas clave de la evaluación, contribuir a los análisis de los datos y ayudar en la elaboración de recomendaciones o planes de acción.

Incluso más infrecuente, según mi definición, es el uso del proceso que abarca la utilización del aprendizaje de las actividades anteriores, a la formación de un equipo de evaluación. Avisar con antelación de que se va a realizar una evaluación y anunciar el tema, las preguntas que se formularán y los métodos que se aplicarán puede afectar a las personas e inducir su utilización. Esto significa que algunos aspectos de la evaluación pueden empezar a utilizarse mucho antes de que los evaluadores sean designados y comiencen con su trabajo. Es más, los usuarios potenciales pueden no ser miembros del equipo de evaluación y quedar fuera del rango de los que sean contactados por el equipo evaluador.

Esto me lleva a mi primera divergencia del patrón estándar 3+1. El uso del proceso no es algo separado y diferente del uso instrumental, conceptual y legitimador, en el sentido de que los procesos de evaluación (y no sólo los productos de evaluación) pueden tener un uso: instrumental, conceptual y legitimador. En mi opinión, sostengo que *todos* los usos pueden referirse tanto a productos como a procesos de evaluación y ambos con un enfoque elegante y fructífero. En consecuencia, el uso instrumental puede ser un tipo genérico de uso para ambos, productos y procesos de evaluación, el uso conceptual otro, y el legitimador un tercero, y así sucesivamente. De este modo, mi enfoque adopta un patrón 3x2 en sustitución del patrón de uso estándar 3+1.

4. Multiplicar seis usos genéricos por dos aspectos de la evaluación equivale a 12 tipos de uso

Mi segunda divergencia con el patrón estándar 3+1 consiste en agregar los usos genéricos táctico, ritual y constitutivo a la predominante tríada compuesta por los usos instrumental, conceptual y legitimador. Y aún más fundamental y esencial es la tercera divergencia, en la que los seis usos se combinan con el uso del producto y el uso del proceso. Esto significa que los seis usos genéricos se presentarán tanto en forma de uso del producto

como de uso del proceso, conduciendo a un patrón 6x2 en lugar de un patrón 3+1.

En 2013-2014, no conocía ningún trabajo sobre el uso que tomara en serio esta última idea, aunque unos años más tarde me di cuenta de que el discurso había cambiado (véanse Højlund 2014 y Alkin/King 2016:575.). La tríada estándar de usos genéricos se aplica ahora en ambas formas, en el proceso y en el producto. No obstante, aún vale la pena la búsqueda sistemática de esta combinación. Cabe destacar que yo pongo en práctica seis usos genéricos, mientras que otros, tres o cuatro usos. La figura 1 ofrece una visión general del resultado mediante la tipología de los usos de la evaluación en doce partes.

A continuación, presentaré la tipología de los usos de la evaluación en doce partes. Los seis tipos genéricos de uso se elaborarán y discutirán tanto en su modalidad de uso del producto como del proceso. Luego, también intentaré responder a las preguntas cinco y seis de la Tabla 1. ¿Será fructífera la nueva combinación de seis usos genéricos por dos aspectos de la evaluación? ¿Son todos los tipos genéricos de uso, aceptables de un modo ético para los evaluadores serios? Comenzaré por el uso instrumental y continuaré con los demás según el orden en que aparecen en la Figura 1.

5. Uso instrumental de la evaluación

En el uso instrumental, los resultados de la evaluación (incluidas las recomendaciones de actuación) se utilizan como una contribución directa para las decisiones inmediatas (cf. Weiss 2013:137) relacionadas con la intervención pública evaluada y su implementación. Los conocimientos de algunas evaluaciones están tan bien fundamentados y son casos tan claros que los responsables de la toma de decisiones los adoptan para la toma de decisiones prácticas. Los resultados de la evaluación pueden utilizarse para poner fin a la intervención, mantenerla, recortarla o ampliarla, ajustar sus actividades, transformar sus instrumentos de políticas públicas o perfeccionar la formación de su personal (cf. Shadish/Cook/Leviton 1991:54; Weiss 1998:23). En el modo instrumental, los resultados finales y provisionales (productos) marcan la diferencia en el sentido de que las decisiones sobre la intervención y su implementación resultan diferentes de lo que habrían sido sin la evaluación.⁵

Figura 1: La tipología de los usos de la evaluación en doce partes

Usos de la evaluación	Instrumental	Productos
		Procesos
	Iluminación	Productos
		Procesos
	Legitimador	Productos
		Procesos
	Táctico	Productos
		Procesos
	Ritual	Productos
		Procesos
	Constitutivo	Productos
		Procesos

Fuente: Vedung 2015:190

⁵Naturalmente, las definiciones del uso instrumental de la evaluación y de la investigación varían en la bibliografía. Me limitaré a dar algunos ejemplos, sin profundizar en el análisis. "Uso instrumental: acción directa y visible basada en los resultados de la evaluación" (Kirkhart 2000:9); "Esto es uso instrumental: tomar decisiones directas sobre el cambio de programas basándose en los resultados de la evaluación" (Shadish/Cook/Leviton 1991:53); o de modo más elusivo: "Los evaluadores valoran el uso directo o instrumental de sus evaluaciones. Por uso directo se entiende el uso documentado y específico de los resultados de la evaluación por parte de los responsables de la toma de decisiones y otros actores sociales" (Rossi/Freeman/Lipsey 1999:432). Peter Dahler-Larsen (2012:20.): "El uso instrumental se refiere a las evaluaciones que se aplican directamente en concordancia con el objetivo puntual de cambiar, perfeccionar, terminar o mantener una actividad". Nutley/Walter/Davies (2007:36): "el uso instrumental se refiere al impacto directo de las investigaciones sobre las decisiones políticas y prácticas. Identifica la influencia de una investigación concreta en la toma de una decisión específica o en la definición de una solución a un problema concreto". Christina Boswell (2009:5 o 30-32) propone una definición mucho más amplia: "Utilizo el término 'instrumental' para abarcar las diversas formas en que el conocimiento influye en los resultados de las políticas públicas, incluidas las funciones 'de iluminación', 'conceptuales' y 'catalizadoras'".

5.1 Uso instrumental del producto: el caso de los resultados experimentales

La evaluación surgió en la década de 1960 como una rama de un amplio movimiento conductista/positivista en las ciencias sociales en general, con el fin de lograr que el sector público fuera más científico, que se basara en el conocimiento y que fuera eficaz, liberándolo así de teorizaciones ideológicas, impresiones inestables, nepotismo y reglas de oro.⁶ En aquel momento, se creía que el denominado modelo de ingeniería era el mejor dispositivo disponible para llevar a cabo esta tarea de la evaluación. En la figura 2 se describen brevemente las características constitutivas del tipo ideal del modelo de ingeniería.

Figura 2: El modelo de ingeniería para el uso instrumental de la evaluación en las políticas públicas (aprox. 1965)

Investigación en Evaluación	Implementación de Políticas Públicas	Administración
	1. Identificación de los problemas sociales	
	2. Fijación de las metas	
	3. Identificación de lagunas de conocimiento sobre los medios para alcanzar los objetivos.	
	4. Comisión para llevar a cabo ensayos científicos sobre los medios.	
5. Investigación científica sobre los medios para alcanzar los objetivos.		
6. Experimento de dos grupos o diseño de ensayos aleatorizados o de tipo quasi experimental (<i>RCTs</i>) ⁷		
7. Conocimiento de métodos o de medios	8. Uso instrumental de los medios de conocimiento para tomar inmediatas decisiones racionales. DECISIÓN	9. Implementación que llevan a productos (<i>outputs</i>), que llevan a impactos o resultados (<i>outcomes</i>). No contemplados en el modelo.

Figura 2: basado en dibujos a mano de E. Vedung

Fuente: Inspiración extraída de Albæk 1988: xx, 23ff. y Naustdalslid/Reitan 1994: 51.

⁶En la ciencia política, el conductismo, el positivismo e incluso el naturalismo son denominaciones para esta corriente de pensamiento que se centra en los comportamientos observables, dejando de lado las intenciones, ideas y motivos humanos, y aspirando a verdades universales similares a leyes como en la física; también se impuso en la política pública y en la evaluación gubernamental (Simon 1976:37). El *conductismo* (Knutsen/Moses 2019:7; Easton 1962, Eulau 1963, Kerlinger 1964, 1966; Charlesworth 1962; Farr 1995; resumen en Vedung 2018:161); *positivismo* (Knutsen/Moses 2019:7); y *naturalismo* (Knutsen/Moses 2019:7 y 31, 44, 94 y 118).

⁷ El término *Two-group experiments (RDTs)* se refiere a experimentos de dos grupos o ensayos aleatorizados (*Randomized Controlled Trials, RCTs* por sus siglas en inglés). Estos son un diseño experimental en el cual los participantes se asignan aleatoriamente a dos grupos: el grupo de tratamiento y el grupo de control. La aleatorización se realiza con el propósito de minimizar los sesgos y asegurar que cualquier diferencia observada entre los dos grupos sea atribuible al tratamiento y no a otros factores. Con mayor detalle:

1. Grupo de Tratamiento (*Treatment Group*): Este grupo recibe la intervención o tratamiento que se está evaluando. Puede ser un nuevo enfoque, una terapia, un medicamento, o cualquier intervención específica que esté siendo estudiada. (Ej. Se les suministra el medicamento)

2. Grupo de Control (*Control Group*): Este grupo no recibe la intervención y se utiliza como punto de referencia para comparar los resultados con el grupo de tratamiento. El objetivo es determinar si la intervención tiene un efecto significativo en comparación con la ausencia de intervención (o la intervención estándar existente). (Ej. No se les suministra el medicamento)

El modelo de ingeniería, un tipo ideal en el sentido weberiano, contiene una doctrina de gobernanza simplificada en seis pasos consecutivos.⁸ En el Paso 1, un fenómeno de la sociedad o de la naturaleza se distingue y se enmarca como un problema, por ejemplo, para investigadores, medios de comunicación, grupos de interés o para un empresario político ocasional. Luego (Paso 2), el problema se incorpora en la agenda política nacional, donde se determinará su tamaño, las fluctuaciones en el tiempo, sus causas, y sus consecuencias en todo el país para facilitar la toma de decisiones sobre si debe ser objeto de intervención pública. En el Paso 3, el sistema político nacional circunscribe el problema⁹, establece los objetivos y encomienda a los académicos que lleven a cabo una investigación rigurosa sobre los medios, ya que los conocimientos disponibles sobre los instrumentos a adoptar en la situación particular en cuestión son de resultado incierto y escasos.

En el Paso 4, los investigadores llevan a cabo la tarea utilizando los mejores diseños de estudios científicos disponibles, es decir, los ensayos controlados aleatorios (*por sus siglas en inglés: Randomized control Trials (RCTs)*). Tras realizar varios *RCTs*, la información encontrada sobre los medios que se consideran más eficaces para conseguir los objetivos establecidos es transmitida a los responsables de la evaluación (el organismo nacional de toma de decisiones, Paso 5). Los responsables de la toma de decisiones aceptan y utilizan estos resultados directa e inmediatamente en la intervención ya preparada, e institucionalizan la intervención mediante una decisión oficial, vinculante para todos los actores de la política nacional (Paso 6). El modelo de ingeniería para el uso instrumental se detiene ahí. Para ilustrar qué partes de un proceso de política pública más completo quedan fuera del modelo, la Figura 2 también contiene la transmisión y la implementación de la maquinaria administrativa que produce el resultado.¹⁰

Se supone que los responsables de la toma de decisiones utilizan directa e inmediatamente los resultados analíticos generados por el tipo ideal, modelo de ingeniería de siete pasos, para influir en el proceso de intervención continua. Este sería el uso instrumental de tipo ideal, que es muy selecto y especial, dado que los resultados de la evaluación utilizados se obtienen de los *RCTs*.

⁸A pesar de la presencia de la palabra ideal en su denominación, un tipo ideal no expresa un ideal en el sentido de algo perfecto o extremadamente favorable por lo que uno debería esforzarse en conseguir. Para Weber, ideal significa de las ideas, idea, el reino de las imágenes. En consecuencia, el modelo de Ingeniería no es una teoría normativa que nos diga cómo debe desarrollarse la interacción entre el organismo encargado de la toma de decisiones, la evaluación académica a través de los *RCTs* y el uso instrumental de los resultados de los *RCTs*. Interpretarlo como una teoría empírica en cierto modo es mejor, porque un tipo ideal se forma a partir de características y elementos de un fenómeno empírico dado. Sin embargo, no se pretende que reúna a todas las características de un caso concreto, sino a algunas, y eso, de una manera muy estilizada, simplificada y estandarizada. Se omiten muchas propiedades y detalles para facilitar la función del tipo como un instrumento para un fin concreto. Un mapa del tamaño de un billete de euro que represente la ciudad de Copenhague mostrando nada más que las líneas del metro y sus estaciones es un tipo ideal muy útil para las personas que quieren utilizar el subte. Podría validarse por su utilidad para este particular propósito. En el presente contexto, el modelo de ingeniería es una herramienta pedagógica para transmitir adecuadamente algunas ideas muy básicas sobre la relación entre un sistema político nacional y una investigación académica sobre los medios de diseño políticos como los *RCTs*.

⁹En raras ocasiones, algunos defensores del uso instrumental de los resultados de la experimentación aleatoria en la formulación de políticas públicas muestran ocasionalmente una tendencia a escribir como si los problemas se hubieran descubierto de forma objetiva en el mundo, no contextualizados y valorados por agentes humanos. Por mi parte, estoy de acuerdo con las siguientes afirmaciones de Becker: "[L]os problemas sociales [...] no son problemas sociales porque estén en su naturaleza ser problemas. Son problemas porque alguien, en algún lugar, los siente y los define como problemas" (1998:162) y Stone: "Formular un problema significa seleccionar algunos aspectos de la realidad percibida y tomarlos como sobresalientes [importantes o visibles] [...] de tal manera que se promueva una definición particular del problema, una interpretación causal, una evaluación moral y/o una recomendación de su tratamiento" (2012: 252).

¹⁰Con sus ventajas y desventajas, la lógica de los *RCTs* como herramienta para la evaluación de políticas públicas (no para la investigación académica básica), se expone en numerosas publicaciones sobre la evaluación; permítanme, en este artículo, hacer unas autorreferencias a Vedung 1997: 170 (2ª ed., próxima publicación en 2022) y Vedung 2013:44.

En un sentido más amplio, tiene su origen en la doctrina de la racionalidad medios-fines que emana de las enseñanzas de Weber. Una elección racional equivale a seleccionar los medios más eficaces para conseguir un fin determinado (objetivo). La racionalidad de objetivos últimos en cadenas de medios-fines no puede determinarse científicamente en su totalidad. En cambio, la ciencia puede identificar qué medios son los más racionales para maximizar la consecución de los objetivos últimos. Siempre y cuando los objetivos últimos sean fijados por organismos ajenos a la comunidad científica y se reconozcan expresamente como subjetivos, la investigación científica podrá producir verdades objetivas e imparciales sobre los medios para alcanzar estos objetivos, si los medios (intervenciones) se prueban cuidadosamente en *RCTs* (racionalidad instrumental).

En un principio, el enfoque del uso instrumental basado en el modelo de ingeniería se consideró interesante, razonable y digno de poner a prueba. Dado que los *RCTs* se han considerado óptimos en las ciencias académicas, ¿por qué no deberían sus resultados también utilizarse directa e inmediatamente en las políticas públicas?

Con el tiempo, el modelo empezó a suscitar críticas. Al parecer, la información basada en los *RCTs* rara vez, o nunca, se utilizaba para la elaboración de políticas públicas. Continuó un gran debate que es demasiado copioso y, ocasionalmente, también profundo para abordarlo aquí. Por razones obvias, me concentro en el tipo de uso instrumental. Para que el argumento a favor del uso instrumental sea lo más sólido posible, asumo que los más máximos responsables oficiales de la toma de decisiones reciben información basada en los *RCTs* de forma oportuna, a un costo razonable y creen en su objetividad y veracidad. Sin embargo, aún tendrán serias dudas sobre su uso directo e inmediato. ¿Por qué?

A continuación se describirá, de manera estilizada y sucinta, la evolución del acrecentamiento de las aprensiones y dudas de los máximos responsables de la toma de decisiones. Las situaciones de decisión en materia de políticas públicas son muy complejas, mucho más intrincadas que lo previsto por los defensores del uso instrumental de la información basada en los *RCTs*. Para empezar, los principales responsables políticos son funcionarios electos en democracia; y deben tener en cuenta la estrategia. Las consideraciones estratégicas son consideraciones de poder. Hay una faceta de poder en la política pública democrática que no se tiene en cuenta, o se tiene, en mucha menor medida, en la ciencia de tipo ideal. La elaboración de políticas públicas y la política en general no son un juego en contra de la naturaleza, sino un juego en contra de otros jugadores. Entre los politólogos y los teóricos de la evaluación, estos otros jugadores son los grupos de interés, que son actores con un interés en su implementación y que afectan o, alternativamente, son o pueden ser afectados por una intervención.

Los responsables de la toma de decisiones deben contemplar qué respuestas suscitarán sus propias acciones, si se toman, en otros grupos de interés y cómo estos conocimientos deberían modificar sus propias acciones. ¿Cómo influirán mis elecciones, entre varias alternativas sustanciales, en las elecciones de otros grupos de interés, y cómo esto afectará a mis actuales decisiones, de modo que tenga que ajustar mis preferencias y mis acciones iniciales?

En primer lugar, la información basada en los *RCTs* sobre los medios puede ir en contra de la ideología de los dirigentes del propio partido. Ante la expectativa de puntos de vista discordantes entre los técnicos y las bases del partido, los máximos responsables de la toma de decisiones vacilarán y evitarán precipitarse sobre el tema. Ya que ellos consideran que los intereses de su propio partido deben estar asegurados antes de proceder. En segundo lugar, los responsables de la toma de decisiones probablemente tendrán que considerar

también las respuestas de otros potenciales grupos de interés: electores, socios de coalición y partidos de oposición parlamentaria. Deben calcular cómo responderán a sus decisiones los medios de comunicación, importantes organizaciones económicas y de la sociedad civil, y reconocidos científicos e investigadores. Todo esto dificulta la utilización directa e inmediata por parte de los máximos responsables de la toma de decisiones incluso contando con información sólida basada en los *RCTs*. Serán necesarios algunos cálculos entre los altos mandos y, algunas deliberaciones (e intrigas) entre los responsables de la toma de decisiones y algunos actores sociales relevantes, antes de que se tomen las decisiones centrales, por decir lo menos.

A pesar de su solidez metodológica, a los ojos de nuestros máximos responsables de la toma de decisiones, la información basada en los *RCTs* es limitada. Los principales responsables de la toma de decisiones necesitan de mucha más información que la que proporcionan los *RCTs*. Los *RCTs* son demasiado limitados para abarcar todos los requerimientos. Los *policy makers* también deben considerar la información sobre la incidencia social del problema sustancial subyacente y los efectos del mismo en los seres humanos y la naturaleza, en la evolución del problema a lo largo del tiempo y en un futuro próximo si no se interviene, y la información sobre las distintas posibilidades para establecer objetivos. También necesitan información sobre los obstáculos y oportunidades de una implementación, en caso de que se tome una decisión política oficial. En todas estas cuestiones, los *RCTs* sobre los medios no son de ninguna ayuda.

Y, por último, los máximos responsables de la toma de decisiones deben considerar la cuestión más amplia de los costos generados por la adopción y la implementación de los medios evaluados científicamente. Además, deben compensar estos costos con los costos contraídos en otras áreas de la política pública y con los recursos disponibles para cubrir todos estos costos. En resumen, no pueden aceptar el supuesto de la comunidad científica, en ocasiones pretencioso, de que el método de los *RCTs* encarna la única forma aceptable de evaluar el sector público.

La combinación de todo esto puede impedir el uso inmediato y directo de los *RCTs* a favor de una recopilación adicional de datos sobre costos, análisis de implementación y opiniones de los grupos de interés, que consume mucho tiempo. También puede ocurrir que se realicen reuniones adicionales para deliberar entre los actores interesados. Por otro lado, la información no experimental puede proporcionar mayores niveles de uso instrumental del producto por encima de una variedad más amplia de problemas políticos (cf. Albæk, 1996, 1995; Boswell, 2009: 29ff.; Nutley/Walter/Davies, 2007; Rossi/Feeman/Lipsey 1999:432; Valovirta, 2002; Weiss, 1979).

Otra salida es centrarse en el uso instrumental del proceso en lugar del uso del producto. La evaluación orientada al uso de Patton es un famoso ejemplo de esta idea alternativa sobre el uso instrumental de las características del proceso.

5.2 Uso instrumental del proceso - Evaluación orientada al uso de Patton

El enfoque alternativo de Patton, denominado evaluación orientada al uso, genera una relación cercana, cara a cara, interactiva, dialógica y conversacional entre los evaluadores y los usuarios destinatarios de la evaluación. Su objetivo es lograr un uso directo y local en la toma de decisiones sobre la intervención evaluada, no se trata de los resultados registrados en un informe preliminar o final, sino de los conocimientos obtenidos a partir de procesos de evaluación continuos, colaborativos, recíprocos, de investigación, de ir y venir y deliberativos que se desarrollan en el territorio. No le preocupan los estudios *RCTs* de dos grupos ni el

requisito de ser científico o académico. Está totalmente a favor del uso del proceso. Cuando un entrevistador le preguntó en 2002, Patton (2002) clasificó el uso del proceso como la número uno, entre las ideas innovadoras en evaluación, de los últimos 10 a 15 años. Uso del proceso

"se produce cuando las personas que participan en una evaluación aprenden del propio proceso de evaluación o realizan cambios en el programa basándose más en el proceso de evaluación que sólo en los resultados de la evaluación. El uso del proceso, por tanto, incluye cambios cognitivos, de actitud y de comportamiento en los individuos, y cambios en el programa u organizativos resultantes, directa o indirectamente, de la participación en el proceso de evaluación y del aprendizaje de pensar de forma evaluativa...". Un ejemplo o prueba del uso del proceso es cuando los participantes de una evaluación más tarde dicen algo como esto: "El impacto en nuestro programa no sólo provino de los resultados, sino también de pasar por el proceso de pensamiento que la evaluación requería" (Patton, 2012: 140ff., también Kirkhart, 2000:10).

La evaluación orientada al uso puede incluir cualquier propósito evaluativo (formativo, de impacto, de desarrollo), cualquier clase de datos (cuantitativos, cualitativos, mixtos), cualquier clase de diseño (naturalista, experimental) y cualquier clase de enfoque (procesos, resultados, impactos, costos, costo-beneficio). La evaluación orientada al uso es un proceso para tomar decisiones sobre estas cuestiones en colaboración con un grupo identificado de usuarios primarios que se centran en los usos que pretenden dar a la evaluación.

En mi interpretación, el uso del proceso de Patton ciertamente incluye conocimientos acerca de los medios para alcanzar los fines establecidos, pero también muchas otras cosas, como objetivos y resultados de la intervención, los problemas de la implementación, los costos, las descripciones de la intervención, las explicaciones, los conceptos y los métodos.

"Uso intencional por usuarios destinatarios" es, en pocas palabras, lo que Patton entiende por utilización (cf. Alkin 1990; Patton 1998:228, 2012:54). Su metodología puede aplicarse al evaluador individual, práctico, y trabajador:

"Mi primera responsabilidad [...] es el uso intencional para los usuarios destinatarios. Y eso es lo que yo llamo utilización primaria. También hay una ética que deberíamos difundir a un público más amplio. Pero normalmente, lo que la gente desconocida hace con eso y cómo lo utiliza no es algo de lo que yo me responsabilice. Creo que es interesante encontrar cómo afecta a la gente su difusión y qué tipo de cosas se aprenden, a qué responden y a qué no, pero no soy el principal responsable de lo que hacen [...]. De lo que sí me responsabilizo es del uso intencional para los usuarios destinatarios" (Patton según Alkin 1990:45-46, 19, 27, 192).

El usuario ideal de Patton es aquel al que principalmente se dirige el evaluador, no cualquier grupo de interés más amplio en torno a la intervención evaluada.

Según Patton, la doctrina del uso instrumental de los resultados de la evaluación *científica* basada en los *RCT* es demasiado limitada para los evaluadores y funcionarios, ya que solamente se centra en los medios para alcanzar los fines establecidos y en los procedimientos no dialógicos, no interactivos y no colaborativos. Patton también critica la noción de uso conceptual (iluminación) por ser demasiado débil, descuidada, poco definida e indirecta. Para arrojar más luz sobre la cuestión del uso, recurriré ahora al segundo tipo genérico: el uso conceptual.

6. Uso conceptual (iluminación)

El uso conceptual acaece cuando alguien directamente involucrado o afectado por la intervención evaluada toma en cuenta los resultados de esa evaluación y aprende algo sin tomar ninguna decisión ni realizar ninguna acción en relación con el programa o con alguna otra intervención similar. El uso conceptual es aprender sin tomar decisiones ni realizar acciones.

Similar a la utilización de la evaluación orientada al uso de Patton, el uso conceptual es una noción amplia y de alguna manera vaga. Sin embargo, está claro que se refiere a un tipo de uso más débil que el uso instrumental que es directo e inmediato para la toma de decisiones y la implementación de acciones (cf. Weiss 1982:292, 303). El uso conceptual implica que la evaluación influye en la forma de pensar de los usuarios potenciales (es decir, "concienciación", "sensibilización" (Rossi et al. 1999: 432-434)). No se toma ni se lleva a cabo ninguna decisión sobre una acción positiva debido a esta influencia en el pensamiento (Weiss 1979:209). Weiss (1982:98) escribe: "La mayoría de las definiciones de uso [evaluación] implican la implementación directa de... los resultados a una decisión. Nuestra definición de utilización se refiere a tener en cuenta los resultados del estudio". Esta idea restrictiva de tener en cuenta es expresada de forma aún más inequívoca por Hendricks: "Nuestras evaluaciones se utilizan cuando 'los resultados y recomendaciones de nuestra evaluación son *consideradas seriamente* por personas en posición de actuar en base a nuestra información si eligen hacerlo'" (citado en Alkin 1990:23). A menudo, el uso conceptual se denomina *iluminación* (cf. Weiss 1981: 20.) o, rara vez, *clarificación* (cf. Kirkhart 2000: 9; Shadish/Cook/Leviton 1991:54).

A diferencia del uso instrumental de tipo ideal del modelo de ingeniería, el uso conceptual también implica el aprendizaje de muchas otras cosas además de los medios para alcanzar los fines determinados. El uso conceptual también abarca cuestiones como las causas y efectos de los problemas de la implementación, las interpretaciones divergentes de los objetivos y las descripciones de las intervenciones, los facilitadores y los obstáculos de la implementación, las opiniones de los destinatarios, el personal operativo y otros grupos de interés y, por último, los costos.

A continuación, desarrollaré la noción conceptual/iluminación más allá del significado básico hasta ahora indicado. Empezaré por el uso del producto y más adelante haré algunas observaciones sobre el uso del proceso.

6.1 Uso conceptual del producto

En el uso conceptual, la consideración de los resultados de la evaluación, conclusiones, recomendaciones para la acción, puede proporcionar conceptos, categorías, etiquetas, marcos, incluso metáforas de experiencias anteriores desorganizadas, como las dificultades de la intervención, las causas sobre las dificultades de la intervención y los efectos de dichas dificultades; pero no se toma ninguna decisión o acción directa e inmediata relacionada con la intervención evaluada (Weiss 1980: 302, 309; 1982: 289-290).

En segundo lugar, a diferencia de los partidarios del uso instrumental directo e inmediato, los defensores del uso conceptual afirman que los resultados no hablan por sí mismos, sino que son interpretados, incluso traducidos y transformados, por el usuario receptor. La mente del usuario no es una pizarra en blanco (*tabula rasa*), que registra exactamente las señales que llegan a través de los sentidos sin procesarlas de ninguna manera. Al contrario, los mensajes se filtran, se clasifican y se etiquetan a través de mapas cognitivos preexistentes, esquemas de codificación, cuadros perceptivos, prismas de ideas y paradigmas. La recepción de nueva información *se enmarca mentalmente* en estas cogniciones del receptor. Los receptores no permiten que los datos sensoriales surjan desde el exterior pasivamente

inalterados; al contrario, los clasifican, transforman y enmarcan de forma activa y selectiva (cf. Boswell 2009:45).

En tercer lugar, el uso conceptual del producto también implica aprender de los resultados generados por la evaluación en conjunto con las conclusiones de *varias evaluaciones*, por ejemplo, recopiladas en revisiones sistemáticas y directrices prácticas. El aprendizaje también puede generarse en combinación con fuentes de conocimiento no evaluativas, como la investigación académica, la experiencia de eficacia comprobada, los entendimientos tácitos, el sentido común, el conocimiento ordinario, la sabiduría profesional, las observaciones personales o incluso los conocimientos derivados del grupo objetivo de la intervención y otros grupos de interés (cf. Weiss 1979 206; 1982:29; Weiss/Bucuvalas 1980:161, 168.).¹¹ Esta visión de que la evaluación no es la única fuente de información implica un fundamento considerablemente más amplio, orientado al sistema y de largo plazo para el uso, que lo expresado al principio de esta sección.

Las tres observaciones sobre el uso conceptual del producto que se acaban de describir sugieren que comprende un aspecto temporal más difuso y menos inmediato. La consolidación paso a paso o el aumento gradual del uso conceptual pueden inducir decisiones y acciones en una perspectiva de largo plazo (Alkin/King 2016:573). Debería investigarse más sobre este tema. Con esta conclusión, dejaré el uso conceptual del producto y me enfocaré ahora en el uso conceptual del proceso.

6.2 Uso conceptual del proceso

La información obtenida mediante las actividades del proceso de evaluación también puede utilizarse para reunir conocimientos (iluminación), sin que vayan acompañados de una acción. O tales aportes, ocasionalmente en combinación con conocimientos procedentes de otras fuentes, pueden conducir a la toma de decisiones y a la adopción de medidas en el largo plazo. Se dan a entender aquí algunos mecanismos del proceso de evaluación que podrían iniciar este tipo de uso del proceso.

Las iniciativas y proyectos locales pueden formar parte de políticas intergubernamentales y multisectoriales más amplias con varios niveles verticales de gobierno. Los programas dirigidos y financiados por la UE son excelentes ejemplos de ello. Al participar el personal y los funcionarios locales, en evaluaciones nacionales o regionales de iniciativas que emanan de los Estados miembros de la UE pueden hacerse una idea mucho más clara y amplia del panorama *vertical* (administrativo) y *horizontal* (geográfico), de las cuales sus actividades particulares locales representan una parte limitada. Así, aprenden algo importante que desconocían antes, aunque no los incite a tomar decisiones y actuar, teniendo en cuenta las circunstancias, en sus propias localidades. O, alternativamente, puede potenciar su moral e influir en sus actividades locales de más largo plazo. Esto no tiene nada que ver con los productos de la evaluación, es decir, los resultados provisionales o finales. Todo esto se debe a los conocimientos obtenidos durante el proceso de evaluación.

En cadenas de implementación complicadas y paralelas, las diversas categorías de personal pueden alimentar una variedad de puntos de vista cognitivos, emotivos y volitivos. Los procesos de evaluación pueden hacer explícitas estas diferencias y concientizar a los afectados. Las personas pueden llegar a comprender mejor las motivaciones de los demás y también a respetar mejor estas diferencias. Esto podría crear percepciones compartidas y

¹¹ El uso conceptual del proceso se refiere a los impactos indirectos de la evaluación en las intervenciones, como políticas públicas, programas y procedimientos. "Este impacto va desde sensibilizar a las personas y grupos sobre los problemas sociales actuales y emergentes, hasta influir en el futuro desarrollo de programas y políticas, contribuyendo a resultados acumulativos o a una serie de evaluaciones" (Rossi et al. 1999:433).

entendimientos, reforzando las actividades de intervención en el largo plazo (cf. Forss/Rebien/Carlsson 2002: 34).

En resumen, el uso conceptual implica tener en cuenta los conocimientos que surgen de los procesos y productos de la evaluación, lo que puede estimular las conversaciones grupales con un fin deliberativo, los argumentos, los contraargumentos, el razonamiento, el aprendizaje y el pensamiento, pero no la toma de decisiones y la acción directa, inmediata o intermedia sobre el programa evaluado o sobre intervenciones similares. El uso conceptual es más integral en su dimensión sustancial (naturaleza de los problemas, rol de las actividades locales en los resultados que abarquen varios sitios o contextos nacionales e internacionales, establecimiento de objetivos, ajuste de objetivos, medios, implementación y clarificación de conceptos), que el uso instrumental, que según el modelo de ingeniería se centra únicamente en los *RCTs* y en el conocimiento científico de los medios. Aunque a menudo es descartado por los partidarios del uso instrumental directo e inmediato, el uso conceptual, deliberativo o no, parece totalmente aceptable (Karlsson 1996, 2001; Valovirta 2002; Vedung 2020). En el discurso evaluativo sobre el uso conceptual también hay una perspectiva amplia, difusa, algo oculta, de implementación a largo plazo, que también parece aceptable.

7 Uso legitimador

En el uso legitimador, los conocimientos de la evaluación se emplean de la manera en que un ebrio utiliza un poste de luz: para apoyarse, no para iluminación. Los resultados se aplican para *respaldar posiciones previamente establecidas*. Entran en escena *después* de que las posturas positivas o negativas sobre determinados temas ya han sido adoptadas y de que los adversarios o aliados parezcan más razonables, respetables y confiables que antes. Las evaluaciones se utilizan para exonerar (blanquear) o denunciar (culpar) políticas, prácticas e instituciones sin cambiarlas. La evaluación sólo proporciona municiones argumentativas adicionales para las batallas discursivas, en las que las alianzas ya están formadas y los frentes ya están establecidos (Boswell 2009: 40 y 61; Albæk 1996:15; Højlund 2014:432).¹²

7.1 Uso legitimador de los productos de la evaluación

Con frecuencia, el uso legitimador del producto es selectivo. Los resultados y recomendaciones que se consideren adecuados para cumplir con los objetivos acordados se toman como evidencia adicional para defender la causa en cuestión, mientras que los resultados que se consideren perjudiciales para ello se ignoran, o se les otorga menos peso argumentativo o se descartan sumariamente. Estas selecciones sesgadas pueden hacerse inteligentemente y deliberadamente.

Reflexionemos sobre dos escenarios específicos. En el escenario A, el consejo de administración de una agencia se encuentra con críticas internas sobre el proceso de cambio organizacional en curso. El consejo está convencido de que los cambios vigentes están justificados, dada la situación actual. Se lleva a cabo una evaluación. Sus resultados suscitan tanto elogios como críticas. Para pulverizar los ataques de los críticos internos, el consejo utiliza una serie de argumentos positivos de la evaluación para reforzar la posición adoptada; los argumentos negativos sólo se mencionan de pasada. Este despliegue de argumentos procedentes de informes de evaluación en apoyo de una postura previamente adoptada sobre determinadas cuestiones ejemplifica el uso legitimador del producto (cf. Boswell 2009:11).

¹² Højlund's (2014: 432) trojka (simbólico, estratégico, legitimador) se incluye en mi uso legitimador.

Consideremos ahora el escenario B. Para aplacar la cobertura negativa de los medios de comunicación, un agente sin influencia en la administración de una evaluación utiliza algunas de sus conclusiones para apoyar su propia postura sobre la intervención. Además, en este caso, la evaluación aporta nuevos argumentos para defender la posición adoptada. Los resultados de la evaluación se emplean como municiones para demostrar el compromiso y debilitar las críticas externas de los actores del entorno (cf. Boswell 2009:11, 24). Esto también ejemplifica el uso legitimador del producto.

7.2 Uso legitimador de los conocimientos del proceso

A veces, los procesos de evaluación también se utilizan para aportar legitimidad. Normalmente, estos usos se producen al principio de los procesos; de hecho, antes de que se reúnan los datos y se elaboren las conclusiones. ¿Por qué ocurre esto?

Según Weiss (1998:25), la comunicación del simple hecho de que se ha iniciado o está en curso una evaluación de rendición de cuentas puede contribuir a deslegitimar la intervención y su gestión ante los ojos de otros actores significativos. Se envía el mensaje de que la intervención está plagada de graves problemas, lo que puede socavar la reputación y legitimidad de la propia intervención. La comunicación empleada de este manera es un ejemplo del uso legitimador del proceso.

¿Debería ser aceptable para los evaluadores el uso legitimador? El discurso sobre la utilización del mismo, en el sector público, está ampliamente desacreditado. Uno de sus escasos defensores es Giandomenico Majone (cf. 1989:31): "Los argumentos justificativos desempeñan un [...] gran [...] papel en la elaboración de políticas. Decidir [...] nunca es suficiente en política. Las decisiones deben legitimarse, aceptarse y llevarse a cabo. Tras el momento de la elección viene el proceso de justificación, explicación y persuasión. Además, los *policymakers* a menudo actúan bajo la presión de acontecimientos externos o según la fuerza de sus convicciones personales. En tales casos, los argumentos son necesarios *una vez tomada la decisión* para proporcionar una base conceptual a la misma, demostrando que es apropiado en el marco de la política existente, para aumentar el consentimiento, descubrir nuevas implicancias y anticipar o responder a las críticas".

Mi respuesta es: depende. Muchas veces, legitimar puede ser defensivo y negativo, pero no siempre. Los resultados pueden legitimar posiciones minoritarias aparentemente extrañas y facilitar su supervivencia, lo que, debido al desarrollo de la sociedad, puede hacer que en el largo plazo lleguen a ser totalmente aceptadas por las grandes mayorías y den lugar a acciones generalmente favorables. La visión popular del calentamiento global de baja a alta aceptación, es un caso de este tipo. En otras situaciones, los resultados pueden proporcionar municiones adicionales a las posiciones existentes que faciliten la generación de una pluralidad de contraargumentos y continuas deliberaciones que, en el largo plazo, contribuirán al uso conceptual. La anticipación, la esperanza y la elección del momento oportuno pueden desempeñar un papel interesante en la legitimación. Si se anticipa algún uso conceptual e instrumental en un futuro distante, la legitimación puede tolerarse en el presente.

8 Uso táctico – Una excusa para la inacción que puede virar hacia otros usos

EL uso táctico afirma que las decisiones no pueden tomarse ahora porque se va a realizar o se está realizando una evaluación; las decisiones deben posponerse hasta que la evaluación presente sus resultados y recomendaciones. El proceso y los resultados de una evaluación próxima o en curso se emplean como excusa para la inacción, para eludir las peticiones de iniciativas y de acción (cf. Weiss 1998:25).

El uso táctico puede emanar de motivos ajenos al pensamiento evaluativo directo. Detrás de las postergaciones puede haber agendas ocultas. Considere la siguiente situación: Estalló una tormenta mediática en relación a los resultados fallidos de un programa. La población local responsabiliza al gobierno, exigiendo que se tomen medidas correctivas inmediatamente. Consternado, el gobierno anuncia la elaboración de una evaluación imparcial sin admitir que su verdadero propósito es desviar la atención hasta que la turbulencia haya disminuido y el programa pueda continuar como antes. Lo que se anuncia abiertamente es que no se tomará ninguna decisión por el momento, ya que se instituirá una evaluación y habrá que esperar que los evaluadores realicen su trabajo. Esto es un uso táctico del proceso.

El uso táctico del producto procede de la misma manera, con la única diferencia de que ahora son los conocimientos de los resultados provisionales y finales, y las recomendaciones las que son utilizadas, y no los procesos de evaluación como tales. La toma de decisiones no puede realizarse ahora porque los resultados de la evaluación se han presentado en un momento inadecuado o no son claros, o están fundamentados de una manera poco convincente, o se consideran deficientes de alguna otra forma. Debido a estas insuficiencias, son necesarios nuevos preparativos adicionales, quizá incluso con una evaluación complementaria. Las propiedades de los resultados de la evaluación se utilizan como una excusa para una nueva postergación. Además, puede haber un cálculo estratégico subyacente.

El uso táctico recuerda al uso legitimador. Incluso en el uso legitimador, el verdadero motivo que impulsa una postura puede estar cuidadosamente camuflado, oculto, suprimido o incluso negado de forma abyecta, por ejemplo, por motivos políticos o personales. El uso legitimador también puede implicar una *acción estratégica*. Sin embargo, también hay una diferencia crucial. En el uso táctico, la evaluación es una excusa para la inacción, mientras que en el uso legitimador, la evaluación se utiliza como un respaldo para posiciones sustanciales ya adoptadas.

La idea del uso táctico ilustra un aspecto vital del uso de la evaluación. Verdaderamente, las intenciones pueden ser latentes, encubiertas o camufladas, como en el uso táctico. También pueden ser manifiestas, comunicarse explícitamente, o darse a conocer al mundo de otra manera, como en el uso intencional de Patton para usuarios destinatarios.

¿El uso táctico es aceptable de un modo ético? Suponiendo que el propósito táctico fuera bastante evidente, ¿deberíamos, como evaluadores serios y moralmente preocupados, aceptar realizar la evaluación cuyos conocimientos sustanciales no sean el objetivo principal de su realización? La respuesta, *No*, parece muy razonable.¹³

Sin embargo, el uso táctico al principio de una evaluación puede convertirse más adelante, en algo más conveniente. Los usos de la evaluación pueden diferir con el tiempo.

¹³Stockmann/Meyer (2013:76): "Sin embargo, este tipo de función 'táctica' difícilmente puede ser aceptada como el verdadero propósito de las evaluaciones. Sería más acertado decir que representa su lado patológico".

Parece que los teóricos de la evaluación han pasado por alto esta idea tan simple. El uso táctico puede ser la justificación aparente al principio de una evaluación; sin embargo, una vez que los resultados se formulan y divulgan, pueden utilizarse de forma instrumental, conceptual o legitimadora. La elección del momento oportuno es importante para los usos de la evaluación. *Anticipando* usos de la evaluación más satisfactorios, cuando los resultados de la evaluación se presenten, el obvio y firme *No* del evaluador al principio, puede convertirse en un *Sí* débil.

9 Uso ritual - Rutinas y procedimientos operativos estándar

Al igual que el uso táctico, el uso ritual suele estar vinculado al proceso de evaluación. Sin embargo, también mostraré que el uso ritual puede ir a la par del producto de la evaluación.

En el uso ritual se hace referencia a algún elemento del *proceso* de evaluación utilizado en un discurso escrito o en una conversación, como una actividad que se anuncia, se lleva a cabo y se emplea de acuerdo con algún procedimiento operativo estándar de su organización patrocinadora. En general, los procedimientos operativos estándar son prácticas que los organismos implementan para garantizar que actividades como la elaboración de presupuestos, la prestación de servicios y la evaluación se lleven a cabo de forma consistente a lo largo del tiempo. La justificación expresa es que las evaluaciones se establecen y se desarrollan de forma rutinaria. Si, una vez finalizada la evaluación, se descubre algo sustancialmente interesante que pueda utilizarse de forma instrumental, conceptual o legitimadora, bien. Sin embargo, ésta no es la justificación inicial de la evaluación. La justificación inicial es *ritualista*, es decir, la evaluación se defiende como una de una serie de actos que siempre se realizan en una situación particular y de la misma manera cada vez. *Simbólico* es otro apelativo menos apropiado para este quinto tipo genérico de uso del proceso de evaluación (cf. Dahler-Larsen 2004:31; 2012:144).

El uso ritual del *producto* de evaluación procede de la misma manera; la única diferencia es que ahora son los resultados y las recomendaciones de la evaluación, y no los procesos, los que se tratan según un procedimiento operativo estándar predeterminado y una rutina establecida. Una prueba del uso ritual de los productos de la evaluación se representan en la siguiente conversación imaginaria, que tiene lugar un par de semanas antes de la fecha límite para presentar el informe final: "Ahora, una vez que la evaluación haya presentado sus resultados finales, los estudiaremos detenidamente, como de costumbre; si no justifican una acción inmediata, estableceremos, como de costumbre, un grupo de estudio para profundizar el asunto y, si es necesario, iniciaremos, también como de costumbre, una nueva evaluación de retroalimentación rápida para complementar los resultados de la evaluación actualmente en curso."

Sin embargo, el uso ritual no es tan vacío como puede parecer. Aunque se hable de una evaluación como un ritual racionalista, simultáneamente puede estar pensada y hablada en secreto o servir como un vehículo oculto, no articulado pero anticipado, para asegurar la legitimidad de los miembros internos de la organización y de los relevantes actores externos, ya sean líderes políticos, intereses organizados o parte del grupo objetivo. Como señaló Boswell (2009:13, 41) "Las organizaciones pueden mejorar su legitimidad adoptando las formas de los estilos racionales de toma de decisiones". El uso ritual argumentativo puede señalar: "adhesión a los códigos culturales de la modernidad reflexiva" (Dahler-Larsen 2004:31), lo que se espera que aumente la legitimidad interna y externa. Hoy en día, todo el mundo evalúa; por tanto, nosotros también debemos evaluar; de lo contrario, se nos considerará atrasados, e irremediamente anticuados (cf. Weiss/Bucuvalas 1980:167). Para nosotros mismos y nuestro entorno, queremos dar a entender que nuestra agencia es una organización racional, moderna, actualizada, que está al día y consciente de los costos, dispuesta a escuchar y aprender de los efectos y efectos secundarios e inclinada a tomar

decisiones bien fundamentadas y meditadas. Por eso evaluamos tanto. Por eso iniciamos esta evaluación en particular. Por eso, naturalmente, siempre examinamos detenidamente y con gran seriedad los resultados de las evaluaciones.

Otros usos esperados además de la legitimación también pueden estar relacionados con el uso ritual. Las evaluaciones pueden utilizarse de forma ritual desde el principio, pero puede preverse que, más adelante, los elementos del proceso de la evaluación se utilicen tácticamente para posponer la toma de decisiones a la luz de agendas ocultas. Y sus productos/resultados, una vez expuestos y perfeccionados, puedan utilizarse de forma instrumental y conceptual. Una vez más, el uso de una misma evaluación puede diferir con el tiempo. Para utilizarla, ¡la elección del momento oportuno es importante!

10 Uso constitutivo – A la sombra del futuro

El sexto y último tipo genérico que se tratará aquí es el uso constitutivo. La palabra constitutivo tiene al menos dos significados distintos:

- 1) Tener poder para instituir, establecer o promulgar (Obtenido de: <https://en.wiktionary.org/wiki/constitute> [08.04.2014])
- 2) Ser parte de la naturaleza más íntima de una persona o cosa; "la proposición de que la libertad y la justicia son elementos *constitutivos* de una sociedad ilustrada" (Obtenido de: <http://www.merriam-webster.com/thesaurus/constitutive?show=-0&t=1396986937>, [08.04.2014])

Trasladado a la evaluación, el uso constitutivo puede referirse al poder de la evaluación para instituir, establecer o promulgar una práctica en el territorio. En segundo lugar, y aquí es decisivo, la evaluación realiza este logro debido a su naturaleza más recóndita. En la bibliografía existente sobre la evaluación, esto se interpreta como el uso antes de que una evaluación produzca resultados; de hecho, incluso antes de que una evaluación haya comenzado su trabajo operativo. El uso constitutivo se interpreta como uso del proceso en un sentido esencial, no como uso de las lecciones aprendidas tras participar de una evaluación y, por supuesto, no como uso de los resultados y recomendaciones de una evaluación (cf. Dahler-Larsen/Krogstrup 2001a: 232).

Dahler-Larsen, pionero y apasionado colaborador de los discursos sobre este sexto tipo genérico, distingue diversas variantes del uso constitutivo del proceso.¹⁴ Una variante bastante ingenua se refiere a las *evaluaciones independientes* llevadas a cabo en la modalidad de rendición de cuentas, no por razones de desarrollo en los que se habían enmarcado los otros cinco tipos genéricos explicados anteriormente. El anuncio de realizar una evaluación de rendición de cuentas junto con la mera transmisión de las áreas temáticas que se van a cubrir, las preguntas que se van a plantear, los criterios y estándares evaluativos que se van a cumplir y los diseños para el análisis de impacto que se van a aplicar para la evaluación inminente pueden proporcionar señales a los subalternos sobre lo que sus superiores esperan de ellos y provocar que respondan cambiando su comportamiento hacia la intervención. El proyecto de evaluación aún no ha producido ningún resultado. El trabajo para el proyecto ni siquiera ha comenzado. Los responsables de la intervención que se va a evaluar esperan lo que pueda venir y ajustan sus actividades de intervención a lo que creen que la evaluación que se aproxima esperará de ellos (cf. Weiss 1982:298). Más tarde, si la evaluación se lleva a cabo, el resultado puede ser la consecución de los objetivos. Si las áreas temáticas, los criterios y las normas transmitidas inicialmente se referían a las actividades fundamentales que debía

¹⁴Cf. Dahler-Larsen 2012:198-320 y 2004: 31; también 2007 y 2011. Enseñar para el examen es un ejemplo al que Dahler-Larsen se refiere a menudo.

emprender la agencia, bien. Si no, las consecuencias de esa anticipación pueden ser perjudiciales (Dahler-Larsen/Krogstrup 2001a: 232).

El uso constitutivo del proceso puede ser involuntario e imprevisto para el funcionario/presentador/director. Pero también puede ser voluntario y previsto. En este último caso, se trata de un doble uso. El funcionario podría estar utilizando la comunicación desde (o próximo) al inicio de una evaluación, consciente del valor del instrumento (utilización 1, uso del proceso por el funcionario) para estimular la gestión de la implementación y al personal local relevante (agentes) para que utilicen este mensaje con el fin de tomar medidas para mejorar la ejecución y los resultados de la intervención (utilización 2, uso del proceso por los agentes) antes de que comience la evaluación propiamente dicha.

Tanto para los funcionarios como para los ejecutores, el *mecanismo de la sombra del futuro* es el motor del uso constitutivo. Está admitido que una actividad evaluativa esperada próximamente proyecte su sombra sobre el presente y desencadene en la actualidad acciones para lograr los objetivos, por los cuales se examinará el valor de la organización en el futuro mediante la evaluación real y, de este modo, se obtendrán elogios y se evitarán críticas.

En segundo lugar, potencialmente muy disfuncional y más complejo, el tipo de uso constitutivo del proceso que se refiere a los sistemas de monitoreo y evaluación (Furubo 2006; Rist/Stame 2006; Hanberger 2013). Desde principios de la década de 1980 y bajo la égida de la doctrina de gobernanza de la gestión por objetivos (también denominada gestión del rendimiento o gestión orientada a resultados), se habían instituido sistemas de monitoreo y evaluación en muchos sectores políticos y programas integrales. Algunos de estos sistemas recurrieron a consentimientos de monitoreo institucionalizados para reunir información de forma regular sobre la implementación, los productos, los resultados y los costos que se consideran interesantes a la luz de las actividades de gobernanza en curso o venideras. Las actividades de gobernanza pueden incluir proyectos por decenas, o por veintenas, o incluso centenares, que deben ir acompañados de una serie equivalente de evaluaciones. En situaciones como éstas, pueden darse varios efectos de procesos constitutivos alternativos.

Veamos un ejemplo ficticio. Un gobierno nacional europeo introduce un programa de subsidios de seis años de duración para mejorar la eficacia de las bibliotecas gestionadas por las autoridades locales. Al mismo tiempo, se instituye un sistema casi permanente de medición del desempeño para supervisar el programa de forma continua. Según los manuales de evaluación, la eficacia debe medirse en la fase de resultados. Un resultado apropiado de una biblioteca sería el aprendizaje y el entretenimiento de los usuarios a partir de la lectura de los libros que toman prestados y las reflexiones y conversaciones concomitantes. Dado que esto es difícil de plasmar y medir, se introduce la productividad como un indicador de rendimiento de la eficacia del resultado, que es más fácil de medir. Se elige como métrica la producción o el número de libros prestados.

Los resultados de la medición se publican cuatro veces al año durante todo el sexenio. Al cabo de un año solamente, varias bibliotecas de todo el país configuran sus actividades a este indicador de libros prestados, porque se mide y se asignan fondos estatales, en función del desarrollo de la productividad. Los responsables de las bibliotecas se dan cuenta de que el indicador es parcialmente deficiente, ya que la producción de las bibliotecas consiste en muchas cosas además del préstamo de libros, como contar con sitios *web*, bases de datos, consultas, periódicos, revistas, exposiciones y conferencias públicas con autores locales (Stroobants/Bouckaert 2014). Inicialmente, también entienden que el valor para los usuarios es haber *leído* los libros, es decir, este resultado, sería un indicador mucho más válido. Pero el número de libros prestados sigue siendo el único indicador. Al cabo de un tiempo, la mayoría de las bibliotecas se adaptan a este indicador hasta tal punto que los libros prestados se convierten en el indicador de buena calidad. Las bibliotecas inician campañas para alentar

a la gente a tomar prestados más libros, principalmente para que sus estadísticas luzcan mejores. Si los usuarios realmente leen los libros es de menor importancia. Esto supondría un tipo gravemente negativo de *desplazamiento de objetivos*, porque el objetivo de una biblioteca pública no es que la gente saque libros; sino que la gente también los lea y aprenda algo o se entretenga con ellos. La gestión y los bibliotecarios trabajan para la siguiente evaluación de su organización, no de acuerdo con la bibliotecología.

Por otra parte, ¿cómo puede darse el uso constitutivo de los productos de evaluación? Durante la redacción de este artículo no he encontrado ningún debate al respecto. Permítanme construir un caso fructífero a partir de diversas formas de razonamiento que he encontrado en mi práctica profesional en mi país natal, Suecia. Consideremos una inspección escolar nacional en un país nórdico, establecida hace dos años por el gobierno para comprobar si los colegios secundarios de todo el país cumplen con los requisitos nacionales. Luego, los responsables del sistema escolar público de un municipio que aún no ha sido evaluado pueden razonar de la siguiente manera: Hemos accedido y estudiado los resultados y recomendaciones de varios informes finales de auditoría disponibles en el sitio *web* del cuerpo de inspectores, cada uno de ellos abarca varias escuelas de varios municipios. Nos hemos percatado de las deficiencias puestas de manifiesto por los inspectores y de las rectificaciones que las escuelas inspeccionadas están obligadas a llevar a cabo. Aprendiendo de ello, y anticipando una próxima auditoría, trabajaremos duro para eliminar los puntos débiles de nuestras propias escuelas. Así, si la inspección se presenta, a las escuelas gestionadas por nuestra municipalidad les irá bien. Recibir puntos de la inspección nos permitirá atraer a más y mejores alumnos y profesores. Y nuestro municipio crecerá y florecerá.

El uso constitutivo bajo la rendición de cuentas del gobierno y la total transparencia de los medios de comunicación puede ser positiva y productiva para la sociedad. Sin embargo, los sistemas de monitoreo y evaluación bienintencionados y más o menos permanentes también tienden a ocultar el desplazamiento de objetivos, fomentando el engaño e incitando comportamientos deshonestos, lo cual es muy perjudicial para los resultados finales. Los diseñadores de sistemas en estos ámbitos deben tener seriamente en cuenta estos riesgos negativos y tomar precauciones. El engaño generalizado destruye la confianza de los ciudadanos y los esfuerzos serios por mejorar los resultados y el bienestar de todos. En algunas circunstancias, sería inteligente evitar por completo tales sistemas. En resumen, del uso constitutivo pueden surgir buenos y malos efectos.

11 ¿Por qué otro artículo sobre el uso de la evaluación?

¿Qué tiene de especial la contribución de este artículo, resumido en la Tipología de los Doce usos de la evaluación? Dado que existen numerosas y excelentes tesis sobre el uso¹⁵, es muy razonable plantearse este interrogante. Permítanme que exponga humildemente mi caso.

La taxonomía fundamental que siguen explicando la mayoría de los expositores del impacto de la evaluación consiste en tratar los usos instrumental, conceptual y legitimador como los tres fundamentales del producto y considerar el uso del proceso como un cuarto tipo de uso, esencialmente diferente de los tres anteriores (véase el epígrafe de Appleton-Dyer et al, al comienzo de este *paper*). En este artículo, he decidido librarme de este patrón 3+1.

¹⁵ Cf. por ejemplo, Alkin 1990, 2013; Alkin/King 2016, 2017; Alkin/Taut 2003; Bjørnholt/Larsen 2014; Boswell 2009; Chouinard/Bourgeois/Amo 2020 (sobre Cousins); Dahler-Larsen 1998, 2001, 2004, 2012, 2014; Hanberger 2013; Højlund 2014; King/Alkin 2018; Kirkhart 2000; Lehtonen 2010; Nutley et al. 2007; Preskill/Zuckerman/Matthews 2003; Rossi/Freeman/Lipsey 1999; Stockmann/Meyer 2013; Valovirta 2002 además del incomparable Patton 1988, 1996, 1998, 2012; y Weiss 1967, 1981, 1988, 1998.

En primer lugar, y muy importante, el uso del proceso no se ha tratado como una cuarta categoría, que no tiene nada que ver con el uso instrumental, conceptual y legitimador. En mi opinión, es mucho más productivo, lógicamente, así como también sustancial, sostener que lo instrumental, conceptual y legitimador está relacionado no sólo con los productos evaluativos, sino también con los procesos evaluativos. Es cierto que esta idea no era completamente original, ni siquiera en 2013-2014, cuando escribí la primera versión de esta contribución (véase Højlund 2014); lo es aún menos ahora en 2020 (véase Alkin/King 2016). No obstante, el propósito sigue mereciendo un seguimiento sistemático.

En segundo lugar, para que la idea de uso del producto/uso del proceso funcione, las he ampliado en dos categorías lógicamente separadas con subcategorías.¹⁶ Un problema es cómo identificar el uso del producto del uso del proceso en lo que respecta a los resultados de la evaluación. Siguiendo una idea que Cousins me señaló en una reunión en Maastricht en 2016 y que Alkin y Taut ya habían propagado en 2003, mi solución es la siguiente. En primer lugar, el uso del producto se refiere a la utilización de los conocimientos evaluativos del informe(s) final(es). Por conocimientos, me refiero a resultados y recomendaciones para la acción. Además, el informe final incluye no sólo el informe(s) final(es) escrito con sus anexos, sino también elementos escritos, orales y audiovisuales que acompañan al informe(s) final(es), como secciones especiales extraídas del informe final y enviadas por separado a audiencias particulares, comunicados de prensa y sesiones informativas orales, combinadas con diapositivas en *PowerPoint*, seminarios, entrevistas y artículos y noticias en los medios de comunicación. Además, y lo que es más importante, la utilización de los resultados y recomendaciones de los informes provisionales y preliminares difundidos durante el proceso de evaluación, transmitidos de forma escrita así como también oral y audiovisualmente, se asignan además al uso del producto.

Aprender de la política evaluada por participar en el proceso de trabajo, como miembro de un equipo de evaluación puede propiciar un uso futuro que se denomina uso del proceso. No ser parte del equipo pero participar de otras maneras, como ser entrevistado por los evaluadores, responder a las encuestas distribuidas por ellos, ayudar a determinar y perfeccionar las preguntas clave de la evaluación, contribuir al desarrollo de criterios de valor o análisis de datos, ayudar en el desarrollo de recomendaciones o planes de acción son además, otras características del proceso que pueden utilizarse.

Asimismo, el anuncio de una próxima evaluación, así como la difusión de información sobre el tema de la evaluación, las preguntas que se formularán y los métodos que se aplicarán, pueden inducir su utilización. Esto significa que un proceso de evaluación puede utilizarse mucho antes de que un equipo de evaluación haya empezado a trabajar, incluso antes de que haya sido designado. Y los usuarios potenciales pueden encontrarse fuera del alcance de los miembros del equipo de la próxima evaluación, así como de los grupos de interés con los que el equipo se pondrá en contacto. Aunque se trate menos en la bibliografía sobre la utilización, el incluido aquí es un tipo de uso relevante en el uso del proceso.

En consecuencia, en línea con Dahler-Larsen pero contrariamente a, por ejemplo, Alkin (cf. 2011:207) y Henry y Mark (cf. 2003:294, y 2004), he argumentado y argumentaré que el uso del proceso incluye mucho más que el uso derivado de la participación en los procedimientos de evaluaciones puntuales; también los actores distantes que no participan en lo más mínimo, de ninguna de las actividades de evaluación en curso pueden utilizar los procesos de evaluación, como los anuncios de próximas evaluaciones y la existencia de

¹⁶ Kirkhart (2000:8.) las identifica como uso orientado a resultados y uso orientado a procesos.

sistemas de monitoreo y evaluación, para mejorar su visión de la intervención en cuestión y la práctica de implementación local de la misma.

En tercer lugar, he añadido el uso táctico y ritual, una diada poco común en la bibliografía de los tres usos genéricos estándar. En cuarto lugar, también he añadido el uso constitutivo como un tipo genérico de uso. El conocimiento de los tres nuevos tipos proporcionará a los evaluadores una comprensión más amplia y profunda de los contextos políticos y administrativos de la evaluación. El uso constitutivo, en particular, sensibilizará a los evaluadores sobre el papel que la anticipación del usuario y el cálculo estratégico, pueden desempeñar, para bien y para mal, una vez que la evaluación se institucionalice en sistemas continuos y semipermanentes.

Por consiguiente, la tabulación cruzada de los seis usos genéricos con los aspectos de evaluación de producto y de proceso da como resultado la tipología de usos de evaluación en doce partes, como se muestra en la figura 2.

En quinto lugar, he intentado profundizar el uso instrumental. El uso instrumental, un concepto bastante difuso, se produce si las ideas inducidas por la evaluación se utilizan directamente en las decisiones y acciones inmediatas relativas a la política en cuestión, es decir, a la intervención evaluada y a su implementación. A título ilustrativo, se explican dos grandes tipos de uso instrumental del producto. Un (tipo) se refiere al uso de los resultados empíricos científicos sobre los medios para alcanzar los objetivos establecidos y producidos externamente por académicos mediante sofisticados experimentos controlados aleatorios de dos grupos. El otro, más flexible desde el punto de vista del diseño metodológico, favorece la realización de entrevistas en profundidad a diversos grupos de interés; y su alcance sustancial es mucho más amplio, abarcando no sólo los medios para alcanzar los objetivos, sino también información sobre, por ejemplo, problemas subyacentes, aclaraciones sobre los objetivos e información sobre la implementación y las organizaciones. Detrás de esta diferencia se esconden diferentes puntos de vista sobre la epistemología: el primero se basa en los métodos y está de acuerdo con la denominada metodología conductista/positivista; y el segundo se basa en la teoría de la intervención y utiliza un enfoque interpretativo/cualitativo de la evaluación.

En sexto lugar, he ampliado y aclarado el uso conceptual. Tradicionalmente, el uso conceptual implica cambios de pensamiento inducidos por la evaluación que no motivan la toma de decisiones ni la adopción de acciones en relación con la intervención evaluada. Para empezar, las evaluaciones pueden utilizarse simplemente para adquirir nuevos conocimientos. Por ejemplo, al participar en evaluaciones de intervenciones locales, que forman parte de iniciativas que se amplían a múltiples sitios y que suponen múltiples niveles jerárquicos intergubernamentales, el personal local puede adquirir una concepción más clara de los amplios panoramas horizontal (múltiples sitios) y vertical (múltiples niveles), de los que sus actividades particulares, restringidas al ámbito local y orientadas al ciudadano, son una parte limitada. Se les informa sobre algo que antes desconocían, aunque eso no los motive a tomar ningún tipo de decisión o acciones.

Además, el uso conceptual es mucho más consciente que el uso instrumental en cuanto al papel del lenguaje. Una primera característica del uso conceptual, no mencionada en el uso instrumental de tipo ideal, es que la evaluación puede proporcionar etiquetas específicas a los actores, conceptos, categorías y metáforas para una temprana organización intelectual de las experiencias. Al mismo tiempo, los mensajes podrían filtrarse, clasificarse y etiquetarse mediante esquemas de codificación del receptor y otros marcos mentales. En el uso conceptual, la mente del usuario no es una *tabula rasa* que registra las señales sensoriales que vienen del exterior sin seleccionarlas, ordenarlas o clasificarlas. El uso conceptual también puede implicar el aprendizaje a partir de nociones generadas por la evaluación en

conjunto con los resultados de varias evaluaciones, por ejemplo, compilados en revisiones sistemáticas resumidas o del saber académico aplicado. Estas tres características implican una ampliación de la perspectiva hacia una visión más holística de la base de uso de la evaluación. De este modo, los cambios de pensamiento pueden dar lugar a decisiones y acciones en el largo plazo.

En séptimo lugar, he utilizado la idea de uso constitutivo de Dahler-Larsen, pero libre de algunos de sus postulados mejor conocidos como uso táctico y especialmente, uso ritual. De algún modo, contrariamente a Dahler-Larsen, también sostengo que puede ser tanto intencionado como no intencionado, y tanto positivo como negativo.

Para terminar, añadiré una octava observación. En consonancia con el discurso académico general sobre el uso, la mayor parte de mi ensayo se centra en evaluaciones únicas e independientes (nivel 1). Sin embargo, el uso también podría relacionarse con los resúmenes de varias evaluaciones sobre el mismo tema, denominados recientemente revisiones sistemáticas, mientras que tradicionalmente se denominaban metaevaluaciones (nivel 2). Además, el uso también podría discutirse en relación con los sistemas de monitoreo y evaluación semipermanentes (nivel 3). Los sistemas de monitoreo y evaluación basados en indicadores producen flujos continuos de información relevante para la evaluación, sobre insumos, procesos, productos y varios resultados consecutivos potenciales de las intervenciones. A partir de los datos de monitoreo sobre costos, procesos y resultados, en combinación con otros datos sobre las causas de las discrepancias con respecto a lo previsto, pueden producirse una serie de evaluaciones de procesos comparables. Además, se supone que se producirán una serie de evaluaciones centradas en los probables efectos de la intervención en las fases de resultados.

Dado que las evaluaciones en los tres niveles se basan en información que difiere en estructura, alcance y tipo de conocimiento, se argumenta que las condiciones previas para el uso también serán diferentes. Por ejemplo, la mera existencia de revisiones sistemáticas o sistemas de monitoreo y evaluación pueden dar lugar a anticipaciones y acciones en los programas pertinentes y entre el personal relevante de los programas.

Sin embargo, la utilización de los resultados y procesos relacionados con los sistemas de evaluación (nivel 3) sólo se ha mencionado de pasada (en relación con el uso constitutivo), pero no se ha tratado de forma sistemática en este artículo. El uso de revisiones sistemáticas y metaevaluaciones (nivel 2) no se ha tratado en absoluto. Esto indica que hay áreas en las que se debe seguir investigando y deliberando sobre el uso de las evaluaciones en la comunidad evaluadora.

Otra mirada muy diferente sobre la cuestión del uso/utilización sería omitir por completo el uso como principal objetivo de la evaluación y sustituirlo por la *utilidad/facilidad de uso*. Como hemos visto, la toma de decisiones prácticas en el sector público, por ejemplo, se ve influenciada por muchos factores además de la evaluación; exigir el uso, y el uso instrumental en particular, de cada evaluación en tales situaciones es simplemente demasiado. *El uso debería sustituirse por la utilidad*. Alcanzar la utilidad sería un requisito justo y razonable para todo evaluador en ejercicio (Rosén et al. 2014).

Agradecimientos

La primera versión publicada de este *paper* se incluyó en la *Nachhaltige Evaluation* en homenaje al profesor Dr. Reinhard Stockmann en su 60 cumpleaños (Hennefeld/Meyer/Silvestrini 2015). En septiembre de 2016, se presentó un informe de alguna manera mejorado en una sesión presidida por Peter Dahler-Larsen en la conferencia de la *European Evaluation Society* en Maastricht. Algunas preguntas y sugerencias de Brad Cousins y Mita Marra motivaron cambios. Los debates en el *Linnaeus Evaluation Research* en Växjö y Kalmar en 2013, 2015, 2016 y 2018, presididos por el profesor Verner Denvall, me suscitaron mayores entendimientos. Ser evaluador del *First Engineer Project (Den första ingenjören)* con Annika Jederström y Staffan Bjurulf también me sirvió de inspiración. De forma intermitente, desde enero hasta noviembre de 2020, he considerado todas las sugerencias y añadido nueva bibliografía sobre el tema. Y desde septiembre de 2020, he ponderado cuidadosamente las valiosas propuestas y críticas de dos revisores anónimos. Hansjörg Gaus ha sido un *Redakteur(editor)* rápido y profesional.

Durante este largo periodo de gestación, muchas personas, además de las ya mencionadas, me habían prestado su apoyo: Magnus Andersson, Merete Watt Boolsen, Thomas Bredgaard, Paulo Calmon, Verner Denvall, Francisco Gaetani, Sebastian Godenhjelm, Annette Gröjer, Nils Hertting, Hanne Kathrine Krogstrup, Kerstin Larsson, Lena Lernå, Sanja Magdalenic, Irene Molina, Kjetil Nordesjö, Luiz Pedone, Gustav J Petersson, Aude Plathner, Mikael Román, Stefan Sjöblom, Jan Sundberg, Kajsa Svanevic, Ulrika Wahlberg, Nicolai Weissert y Tage Vedung. A todos los mencionados, extendiendo mi más sincera gratitud. Naturalmente, toda la responsabilidad del texto final recae exclusivamente en mí.

Edición de idioma: Nancy Aaen, experta en inglés, Roskilde.

Bibliografía

- Albæk, Erik (1988): *Fra Sandhed Til Information: Evalueringsforskning i USA - Før Og Nu*. København: Akademisk Forlag.
- Albæk, Erik, 1995, "Between Knowledge and Power: Utilization of Social Science in Public Policy Making," in *Policy Sciences*, 28, 79–100.
- Albæk, Erik (1996): Why All this Evaluation: Theoretical Notes and Empirical Observations on the Functions and Growth of Evaluation, With Denmark as an Illustrative Case. In: *Canadian Journal of Program Evaluation*, 11 (2), pp. 1-34.
- Alkin, Marvin C. (1990): *Debates on Evaluation*. London: Sage.
- Alkin, Marvin C. (2011): *Evaluation Essentials: From A to Z*. London: Guilford Press.
- Alkin, Marvin C. (Ed.) (2013): *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. Los Angeles: Sage.
- Alkin, Marvin C. & King, Jean A. (2016): The Historical Development of Evaluation Use. In: *American Journal of Evaluation*, 37 (4), pp. 568-579.
- Alkin, Marvin C. & King, Jean A. (2017): Definitions of Evaluation Use and Misuse, Evaluation Influence, and Factors Affecting Use. In: *American Journal of Evaluation*, 38 (3), pp. 434-45.
- Alkin, Marvin C./Taut, Sandy M. (2003): Unbundling Evaluation Use. In: *Studies in Educational Evaluation*, 29, pp. 1-12.
- Appleton-Dyer, Sarah/Clinton, Janet/Carswell, Peter/Mc Neill, Rob (2011, August): What Evaluation Attributes, Stakeholder Characteristics & Contextual Factors Are Important For Evaluation Influence? Sydney: AES Conference, August 2011. (Retrieved from: <https://slideplayer.com/slide/9859342/> [01.10.2020])
- Becker, Howard S. (1998): *Tricks of the Trade: How to Think About Your Research While You're Doing It*. London: University of Chicago Press.
- Bjørnholt, Bente/Larsen, Flemming (2014): The Social Context of Evaluation. In: *Evaluation*, 20 (4), pp. 400-411.
- Boswell, Christina (2009): *The Political Uses of Expert Knowledge: Immigration Policy and Social Research*. Cambridge: Cambridge University Press.
- Charlesworth, James C. (1962): *The Limits of Behavioralism in Political Science: A Symposium Sponsored by the American Academy of Political and Social Science*. Philadelphia: American Academy of Political and Social Science.
- Chouinard, Jill Anne/Bourgeois, Isabelle/Amo, Courtney (Eds.) (2020): *Growing the Knowledge Base in Evaluation: The Contributions of J. Bradley Cousins*. Charlotte: Information Age Publishing.
- Dahler-Larsen, Peter/Krogstrup, Hanne K. (2001a): Evalueringens Konstitutive Virkninger. In: Dahler-Larsen, Peter/Krogstrup, Hanne K. (Eds.): *Tendenser i evaluering*. Odense: Syddansk Universitetsforlag, pp. 232-245.
- Dahler-Larsen, Peter/Larsen, Flemming (2001b): Evalueringsanvendelse: Historien Om Et Begreb, Der Udvider Sig. In: Dahler-Larsen, Peter/Krogstrup, Hanne K. (Eds.): *Tendenser i Evaluering*. Odense: Syddansk Universitetsforlag, pp. ?.
- Dahler-Larsen, Peter/Krogstrup, Hanne K. (2001c): *Nye Veje i Evaluering*. Odense: Syddansk Universitetsforlag.
- Dahler-Larsen, Peter (2004): *Evaluering Og Magt*. Århus: Aarhus universitet, Institut for statskundskab, Magtudredningen.
- Dahler-Larsen, Peter (2007): Constitutive Effects of Performance Indicator Systems. In: Kushner, Saville/Norris Nigel (Eds.): *Dilemmas of Engagement: Evaluation and the New Public Management*. Amsterdam: Heinemann, pp. 17-36.
- Dahler-Larsen, Peter (2011): Constitutive Effects as a Social Accomplishment: A Qualitative Study of the Political in Testing. In: *Education Inquiry*, 3 (2), pp. 171-186.
- Dahler-Larsen, Peter (2012): *The Evaluation Society*. Stanford: Stanford University Press.
- Easton, David (1962): The Meaning of Behavioralism in Political Science. In: Charlesworth, James C.: *The Limits of Behavioralism in Political Science*. Philadelphia: American

- Academy of Political and Social Science, pp. ?.
- Eulau, Heinz (1963): *The Behavioral Persuasion in Politics*. New York: Random House.
- Farr, James (1995): *Remembering the Revolution: Behavioralism in American Political Science*. In: Farr, J./Dryzek, John S./Leonard, Stephen T. (Eds.): *Political Science in History: Research Programs and Political Traditions*. Cambridge: Cambridge University Press.
- Forss, Kim/Rebien, Claus C./Carlsson, Jerker (2002): Process Use of Evaluations Types of Use that Precede Lessons Learned and Feedback. In: *Evaluation*, 8 (1), pp. 29-45.
- Furubo, Jan-Eric (2006): Why Evaluations Sometimes Can't Be Used – And Why They Shouldn't. In: Rist, Ray C./Stame, Nicoletta (Eds.): *From Studies to Streams: Managing Evaluative Systems*. New Brunswick: Transaction, pp. 147-165.
- Hanberger, Anders (2013): Framework for Exploring the Interplay of Governance and Evaluation. In: *Scandinavian Journal of Public Administration (special issue on evaluation)*, 16 (3), pp. 9-27.
- Hennefeld, Vera/Meyer, Wolfgang/Silvestrini, Stefan (Eds.): *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann*. Münster, New York: Waxmann.
- Henry, Gary T./Mark, Melvin M. (2003): Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. In: *American Journal of Evaluation*, 24 (3), pp. 293-314.
- Højlund, Steven (2014): Evaluation Use in Evaluation Systems: the Case of the European Commission. In: *Evaluation*, 20 (4), pp. 428-446.
- Karlsson, Ove (1996): A Critical Dialogue in Evaluation: How can the Interaction Between Politics and Evaluation Be Tackled?. In: *Evaluation*, 2 (4), pp. 405-416.
- Karlsson, Ove (2001): Critical Dialogue: Its Value and Meaning. In: *Evaluation*, 7 (2), pp. 211-227.
- Kerlinger, Fred N. (1964, 1966): *Foundations of Behavioral Research: Educational and Psychological Inquiry*. New York: Holt, Rinehart & Winston.
- King, Jean A./Alkin, Marvin C. (2018): The Centrality of Use: Theories of Evaluation Use and Influence and Thoughts on the First 50 Years of Use Research. In: *American Journal of Evaluation*, 40 (3), pp. 431-458.
- Kirkhart, Karen E. (2000): Reconceptualizing Evaluation Use: An Integrated Theory of Influence. In: *New Directions for Evaluation*, 88, pp. 5-23.
- Knutsen, Torbjørn L./Moses, Johnathon W. (2019): *Ways of Knowing; Competing Methodologies in Social and Political Research*. London: Red Globe Press.
- Lehtonen, Markku (2010, April): *IEA Energy Policy Reviews: Transitions Towards Sustainable Energy?* Paper presented at the workshop Peer Reviews as an Instrument of Global Governance, organized by the Universities of Bremen and Maastricht, Bremen, 9-10 April 2010.
- Majone, Giandomenico (1989): *Evidence, Argument and Persuasion in the Policy Process*. London: Yale University Press.
- Mark, Melvin M./Henry, Gary T. (2004): The Mechanisms and Outcomes of Evaluation Influence. In: *Evaluation*, 10 (1), pp. 35-57.
- Naustdalslid, Jon/Reitan, Marit. (1994): *Kunnskap og styring: Om forskningens rolle i politikk og forvaltning*. Oslo: Tano.
- Nutley, Sandra M./Walter, Isabel/Davies, Hew T. O. (2007): *Using Evidence: How Research Can Inform Public Services*. Bristol: Policy Press.
- Patton, Michael Q. (1997): *Utilization-focused Evaluation: The New Century Text*. Thousand Oaks: Sage.
- Patton, M. Q. (1998): Discovering Process Use. *Evaluation*, 4 (2), pp. 225-233.
- Patton, M. Q. (2002): Interview. *The Evaluation Exchange: A Periodical on Emerging Strategies in Evaluation*, 8 (1).
- Patton, M. Q. (2012): *Essentials of Utilization-Focused Evaluation*. London: Sage.
- Preskill, H./Zuckerman, Barbra/Matthews, Bonya (2003): An Exploratory Study of Process Use: Findings and implications for future research. *American Journal of Evaluation*, 24 (4), pp. 423-442.

- Rist, Ray C./Stame, Nicoletta (2006): *From Studies to Streams: Managing Evaluative Systems*. New Brunswick: Transaction.
- Rosén, Kajsa/Wolf-Watz, Olov/Vedung, Evert (2014): *Usefulness and Utilization*. Uppsala: Evert Vedung's private archive, mimeo.
- Rossi, Peter H./Freeman, H. E./Lipsey, M. W. (1999): *Evaluation: A Systematic Approach*. London: Sage.
- Sartori, Giovanni (2009a): *The Tower of Babel*. In: Collier, David/Gerring, John (Eds.): *Concepts and Method in Social Science: The Tradition of Giovanni Sartori*. London: Routledge, pp. 61-96.
- Sartori, Giovanni (2009b) *Guidelines for Concept Analysis*. In: Collier, David/Gerring, John (Eds.): *Concepts and Method in Social Science: The Tradition of Giovanni Sartori*. London: Routledge, pp. 97-150.
- Scriven, Michael (2013): *Conceptual Revolutions in Evaluation: Past, Present, and Future*. In: Alkin, Marvin C. (Ed.): *Evaluation Roots: A Wider Perspective on Theorists' Views and Influences*. London: Sage, pp. 167-179.
- Shadish, William R. Jr/Cook, Thomas D./Leviton, Laura C. (1991): *Foundations of Program Evaluation: Theories of Practice*. London: Sage.
- Stockmann, Reinhard/Meyer, Wolfgang (2013): *Functions, Methods, and Concepts in Evaluation*. Houndmills: Palgrave Macmillan.
- Stone, Deborah (2012): *The Policy Paradox: The Art of Political Decision Making*. London: Norton.
- Stroobants, Jesse/Bouckaert, Geert (2014): *Benchmarking Local Public Libraries Using Non-Parametric Frontier Methods: A Case Study of Flanders*. <http://dx.doi.org/10.1016/j.lisr.2014.06.002> [05.12.2014].
- Valovirta, Ville (2002): *Evaluation Utilization as Argumentation*. In: *Evaluation*, 8 (1), pp. 60-80.
- Vedung, Evert (1997, 2017): *Public Policy and Program Evaluation*. New York: Transaction 1997, Taylor & Francis, Routledge 2017.
- Vedung, Evert (1999): *Evaluation im öffentlichen Sektor*. Wien, Köln, Weimar, New York: Böhlau Verlag.
- Vedung, Evert (2013): *Side Effects, Perverse Effects and Other Strange Effects of Public Interventions*. In: Svensson, Lennart/Brulin, Göran/ Jansson, Sven/Sjöberg, Karin (Eds.): *Capturing Effects of Projects and Programmes*. Lund: Studentlitteratur, pp. 35-62.
- Vedung, Evert (2015): *Six Uses of Evaluation*. In Hennefeld, Vera/Meyer, Wolfgang/Silvestrini, Stefan (Eds.): *Nachhaltige Evaluation? Auftragsforschung Zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann*. Münster, New York: Waxmann, pp. 187-210.
- Vedung, Evert (2018): *Idéanalys Från Ankdam Till Universums Mitt [Ideational Analysis from Duck Pool to Center of the Universe]*. In: *Statsvetenskaplig Tidskrift*, 120 (2), pp. 155-195.
- Vedung, Evert (2020): *Policy Evaluation*. In: Berg-Schlosser, Dirk/Morlino, Leonardo/Badie, Bertrand (Eds.): *Handbook of Political Science - A Global Perspective*. London: Sage.
- Weiss, Carol H. (1979): *The Many Meanings of Research Utilization*. *Public Administration Review*, 39, pp. 426-431.
- Weiss, Carol H. (1980): *Knowledge Creep and Decision Accretion*. *Knowledge: Creation, Diffusion, Utilization*, 1 (3), pp. 381-404.
- Weiss, Carol H. (1981): *Measuring the Use of Evaluation*. In: Ciarlo, James A. (Ed.): *Utilizing Evaluation: Concepts and Measurement Techniques*. Beverly Hills: Sage, pp. 17-33.
- Weiss, Carol H. (1982): *Policy Research in the Context of Diffuse Decision-Making*. *Social Science Research and Public Policy-Making*. In: Kallen, Denis B. P./Kosse, G. B./Wagenaar, Hendrik C./Kloprogge, J. J. J./Vorbeck, M. (Eds.): *Social Science and Policy-Making: A Reappraisal*. Windsor, Berks: NFER-Nelson, pp. 288-314.
- Weiss, Carol H. (1988): *If Program Decisions Hinged Only on Information: A Response to Patton*. *Evaluation Practice*, 9, pp. 15-28.

- Weiss, Carol H. (1998): Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation*, 19 (1), pp. 21-33.
- Weiss, Carol H. (2013): Rooting for Evaluation: Digging Into Beliefs. In: Alkin, Marvin C. (Ed.): *Evaluation Roots: A Wider Perspective on Theorists' Views and Influences*. London: Sage, pp. 130-143.
- Weiss, Carol H./Bucuvalas, Michael J. (1980): *Social Science Research and Decision-Making*. New York: Columbia University Press.

