

Razonamiento evaluativo

E. Jane Davidson

CENTRO DE INVESTIGACIONES INNOCENTI DE UNICEF

El Centro de Investigaciones Innocenti es la oficina de UNICEF especializada en investigación científica. El objetivo principal del Centro de Investigaciones es mejorar la comprensión internacional de una serie de cuestiones relacionadas con los derechos de la infancia, a fin de facilitar la plena aplicación de la Convención sobre los Derechos del Niño en todo el mundo. El centro tiene el cometido de establecer un marco integral de investigación y conocimiento dentro de la organización para brindar apoyo a los programas y políticas mundiales de UNICEF, y trabaja con los asociados para formular políticas con base empírica en favor de la infancia. Las publicaciones elaboradas por el centro contribuyen al debate global sobre la infancia y los derechos del niño e incluyen una amplia gama de opiniones.

Las opiniones expresadas corresponden a los autores o editores y se publican para estimular un mayor diálogo sobre métodos de análisis de impacto. Esta publicación no refleja necesariamente las políticas o perspectivas de UNICEF.

SINTESIS METODOLOGICAS DEL CENTRO DE INVESTIGACIONES

Las síntesis metodológicas del Centro de Investigaciones de UNICEF pretenden compartir prácticas de investigación, métodos, diseños y recomendaciones de reconocidos investigadores y analistas. La audiencia a la que van dirigidas es principalmente el personal de UNICEF que lleve a cabo, encargue o interprete los resultados de investigación y análisis para la toma de decisiones sobre programas, políticas y actividades de sensibilización.

Esta síntesis metodológica ha seguido un proceso de revisión por pares interna.

El texto no ha sido editado de acuerdo con los estándares de publicación oficiales y UNICEF declina toda responsabilidad por posibles errores.

Se permite la reproducción de cualquier parte de la presente publicación siempre que se incluya referencia a la presente. Si se desea utilizar una parte sustancial o la totalidad de la publicación dirijan su solicitud al Departamento de Comunicación en la dirección de correo electrónico: florence@unicef.org

Para consultas o descargas, pueden encontrar estas síntesis metodológicas en <http://www.unicef-irc.org/KM/IE/>

Recomendamos la siguiente cita para cualquier referencia al presente documento:

Davidson, E. J. (2014). Razonamiento evaluativo, *Síntesis metodológicas: evaluación de impacto n.º 4*, Centro de Investigaciones de UNICEF, Florencia.

Agradecimientos: Varios autores han proporcionado orientación en la preparación de esta síntesis. El autor y el Centro de Investigaciones de UNICEF desean agradecer a todos aquellos que han participado en la preparación de la presente publicación, especialmente a:

Por su contribución: Simon Hearn, Patricia Rogers, Jessica Sinclair Taylor

Por su revisión: Nikola Balvin, Karin Heissler, Debra Jackson

© Fondo de las Naciones Unidas para la Infancia (UNICEF), septiembre de 2014

Centro de Investigaciones Innocenti de UNICEF

Piazza SS. Annunziata, 12

50122 Florencia (Italia)

Tel.: (+39) 055 20 330

Fax: (+39) 055 2033 220

florence@unicef.org

www.unicef-irc.org

1. RAZONAMIENTO EVALUATIVO: BREVE DESCRIPCIÓN

Una evaluación responde, por definición, a una serie de preguntas *evaluativas*, es decir, preguntas sobre la calidad y el valor. Por ese motivo, las evaluaciones son mucho más útiles y pertinentes que la mera medición de indicadores o los resúmenes de observaciones y narraciones.

Los responsables de las decisiones recurren con frecuencia a ellas cuando tratan de dilucidar cómo pueden aprovechar los puntos fuertes y hacer frente a los puntos débiles. Para ello, no solo han de saber cuáles son esos puntos fuertes y débiles, sino también cuáles son más importantes o graves y cuál es el nivel de desempeño del programa o la política correspondiente. Por ejemplo, no basta con que conozcan la medida en que ha cambiado un determinado resultado¹; también deben conocer la *calidad* y *el valor* de dicho resultado.

Para responder a las [preguntas evaluativas](#)², hay que definir en primer lugar qué significan los términos «calidad» y «valor»; posteriormente, deben recogerse las pruebas pertinentes. La calidad se refiere al nivel de idoneidad de algo; el valor se refiere a su idoneidad con respecto a una situación concreta y tiene en cuenta especialmente los recursos que se destinaron a su producción y las necesidades a las que supuestamente hizo frente. En necesario llevar a cabo un razonamiento evaluativo que sintetice esos elementos con miras a formular respuestas justificables (es decir, debidamente razonadas y basadas en pruebas) a las preguntas evaluativas.

El razonamiento evaluativo es un pilar fundamental de la evaluación: se emplea a lo largo de todo el proceso para sintetizar la información necesaria y extraer conclusiones. Esto se consigue de dos maneras, mediante la combinación de:

- pruebas sobre el desempeño en una dimensión particular, que se interpretan en relación con las definiciones del grado de idoneidad para generar una calificación del desempeño en esa dimensión; y
- calificaciones del desempeño en varias dimensiones que dan lugar a una conclusión general sobre su grado de idoneidad en un determinado emplazamiento, proyecto, programa, política u otro «evaluando» (término genérico que designa el objeto de la evaluación).

El razonamiento evaluativo es un requisito de *todas* las evaluaciones —[formativas](#), [sumativas](#) o [evolutivas](#); de [procesos](#), [resultados](#) o [impacto](#); a través de métodos [cualitativos](#), [cuantitativos](#) o [mixtos](#); y [primarias](#) o [secundarias](#)—. Es también obligatorio en *todas* las opciones de diseño.

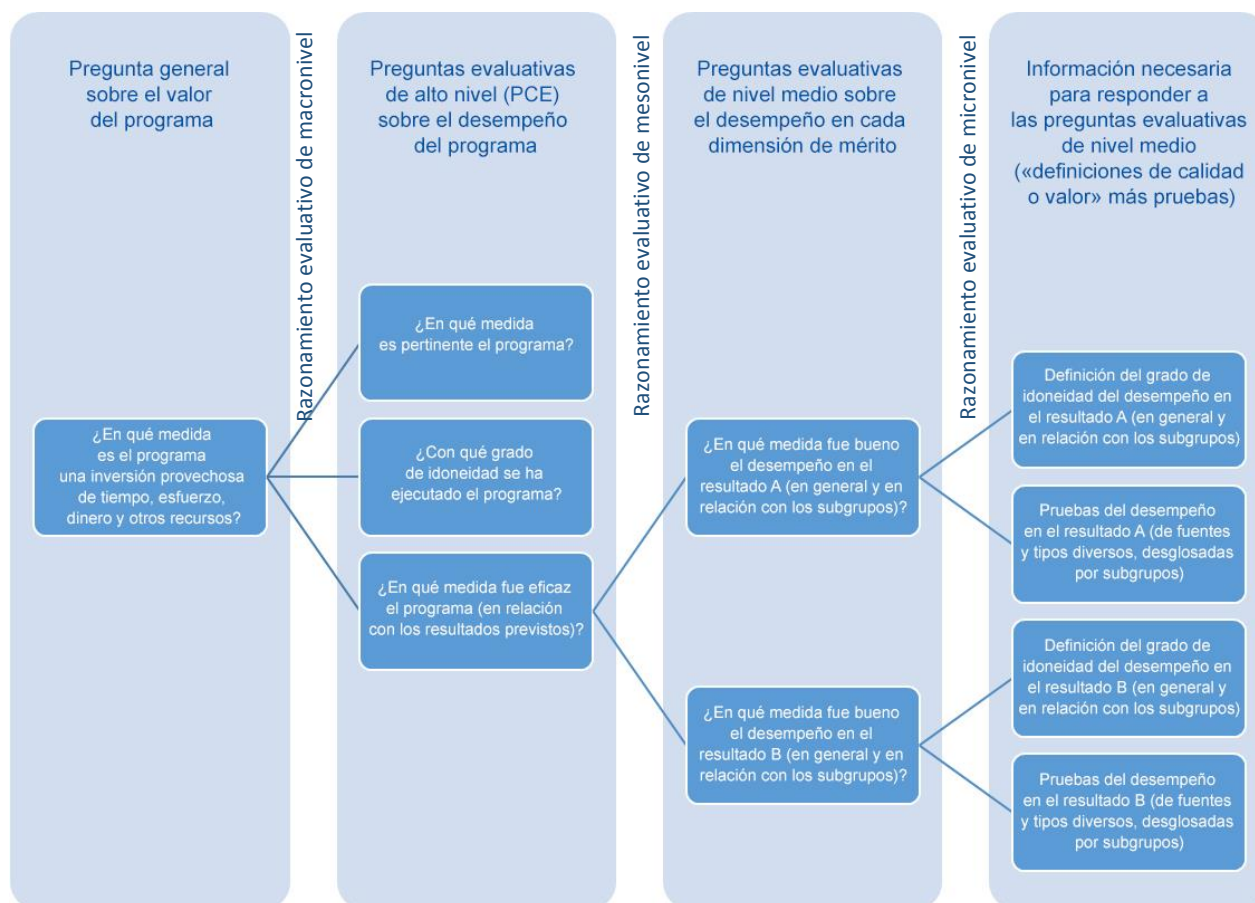
El gráfico 1 muestra cómo se determina el valor general de un programa o una política, al responder a un conjunto de [preguntas clave de evaluación](#) de alto nivel (PCE; véase también la Síntesis n.º 3 (Criterios de evaluación)). Una evaluación debe constar de un conjunto limitado de preguntas de alto nivel sobre el desempeño general. Cada una de esas PCE ha de dividirse posteriormente en preguntas más pormenorizadas relativas al desempeño en dimensiones de mérito específicas (es decir, la PCE se aplica a través de una serie de preguntas de nivel medio (mesonivel), como se aprecia en el gráfico 1, de izquierda a derecha) e incluso en preguntas de nivel inferior (micronivel) (no aparecen en el gráfico 1).

El razonamiento evaluativo es el proceso consistente en sintetizar las respuestas a las preguntas de nivel medio e inferior en juicios justificables que responden de manera directa a las preguntas de alto nivel. Todas las evaluaciones requieren un razonamiento evaluativo de micronivel y mesonivel (es decir, que se sintetizan las respuestas de derecha a izquierda en el gráfico 1); no todas lo exigen a alto nivel.

¹ Un «resultado» es *algo que sucede* a las personas, las comunidades, el medio ambiente o la economía propiciado, al menos parcialmente, por un programa o una política. Puede ser positivo o negativo, previsto o imprevisto, un cambio que tiene lugar o un cambio cuya realización se impide. En el contexto de una evaluación, el término «resultado» implica que el programa o la política en cuestión contribuyeron a dicho cambio, ayudaron a provocarlo o lo catalizaron (aceleraron).

² En la Síntesis n.º 1 (Sinopsis de la evaluación de impacto) se ofrece una panorámica de los distintos tipos de preguntas (descriptivas, causales, evaluativas) que las evaluaciones pueden plantearse.

Gráfico 1. Niveles de razonamiento evaluativo que determinan el valor general de una intervención



El primer paso del razonamiento evaluativo ha de ser la identificación de los [criterios de evaluación](#) (véase la Síntesis n.º 3 (Criterios de evaluación)) y de las pruebas a partir de las cuales se valorará el desempeño; también es preciso determinar cómo se sintetizarán y usarán para generar respuestas debidamente razonadas y justificadas a cada una de las preguntas evaluativas.

El cuadro 1 muestra un ejemplo extraído de la evaluación de un programa de liderazgo educativo, cuyo resumen explica las conclusiones clave de manera expresamente evaluativa. Incluye una declaración inequívoca sobre la calidad del programa e impresiones sobre sus deficiencias principales, así como una destacable valoración de su gravedad. Se ofrece asimismo una breve sinopsis de los fundamentos de tales conclusiones, y se remite al lector a un desglose claro de las constataciones que se brinda en la tabla 1 a través de un gráfico de barras de fácil interpretación con *calificaciones evaluativas explícitas* («excelente», «adecuado», etc.) para cada criterio. De ese modo, se pone de manifiesto el grado de idoneidad del desempeño en cada aspecto.

Cuadro 1. Ejemplo de explicación evaluativa explícita de las constataciones clave

Constataciones clave del estudio

En general, el examen constató que el programa de orientación para directores novatos (FTPIP) ofreció un programa de calidad a sus destinatarios. Aunque algunos aspectos del programa requieren ajustes en aras de una mayor eficacia, el examen no recomendó cambios importantes. Su conclusión es que el FTPIP constituyó una plataforma adecuada de orientación para que los

directores novatos de centros de primaria y secundaria, y (en menor medida) de las escuelas kura kaupapa asuman el liderazgo educativo en sus escuelas.

En mi opinión, el programa para directores novatos fue una iniciativa estupenda y un dinero bien invertido. Me sentí respaldado y siempre pude efectuar las consultas necesarias. Los tutores y los miembros del PLG (grupo de aprendizaje profesional) nos transmitieron sus experiencias y brindaron asesoramiento de forma oportuna. Las sesiones de cada uno de los cursos residenciales fueron tremendamente acertadas y oportunas. Gracias por este fantástico programa. En él se combinaron lecciones contextuales y de la vida real que pudimos compartir con personas y directores experimentados. (Participante en el programa de orientación para directores novatos)

La tabla 1 a continuación ofrece una instantánea en la que se identifican los criterios de evaluación en que los directores novatos, los interesados y los líderes del sector observaron un mejor y peor desempeño.

Por lo general, los directores novatos consideraron que en su mayoría (con algunas excepciones) estaban mejores preparados que antes del curso para asumir el liderazgo en el plano pedagógico y aplicar ese aprendizaje en sus centros escolares.

Fuente: Reproducido bajo una licencia Creative Commons de Oakden, Judy, *Evaluation rubrics: how to ensure transparent and clear assessment that respects diverse lines of evidence*, BetterEvaluation, 2013, p. 12. Véase <http://betterevaluation.org/resource/example/rubrics-oakden>.

Tabla 1. Ejemplo de un desglose claro de constataciones con calificaciones evaluativas explícitas (en relación con el cuadro 1)

Criterios de evaluación		Calificaciones				
		Deficiente	Adecuada	Buena	Muy buena	Excelente
Participación en el programa de capacitación profesional y reconocimiento como «líderes pedagógicos» en sus centros	Calificación general					
	Se establecen de forma clara y adecuada los objetivos de desarrollo profesional de los directores novatos					
	Se establece una buena relación de trabajo entre los directores novatos y los tutores, que les brindan asesoramiento y consejo profesional					
	Se crean redes de apoyo entre pares entre los directores novatos y sus homólogos					
	Los directores novatos se implican en un aprendizaje reflexivo sobre la asunción del liderazgo pedagógico en sus centros					

Los directores novatos entienden la importancia de asumir el liderazgo pedagógico y contar con estrategias claras a tal efecto					
Los directores novatos señalan que saben cómo recopilar, analizar y aplicar los datos para respaldar el aprendizaje de sus alumnos					
Se dispone de pruebas acerca de la atención que prestan los directores novatos a la equidad para los maoríes					
El respaldo a los directores novatos se coordina adecuadamente (especialmente cuando trabajan a la vez con varios grupos de apoyo)					

Fuente: *Ibid.*

Puntos principales

1. Se recurre al razonamiento evaluativo a lo largo de la evaluación para combinar sistemáticamente las pruebas con las definiciones de «calidad» y «valor», y llegar así a conclusiones justificables sobre el desempeño.
2. El razonamiento evaluativo es obligatorio en todas las evaluaciones, con independencia de su propósito, método o diseño.
3. El razonamiento evaluativo es lo contrario de la «operativización». Al operativizar una pregunta evaluativa, esta se desglosa a fin de descubrir qué pruebas es necesario reunir y qué definiciones de calidad y valor se aplicarán. El proceso de razonamiento evaluativo consiste en reagrupar esos elementos para responder a la pregunta evaluativa.

2. ¿CUÁNDO PROCEDE EMPLEAR ESTE MÉTODO?

Todas las evaluaciones requieren un razonamiento evaluativo. Los [términos de referencia](#) de una evaluación deben indicar de forma expresa cómo se llevará a cabo el proceso de razonamiento evaluativo. Al establecer con claridad las expectativas, los directivos contribuyen a que la evaluación brinde respuestas evaluativas directas y explícitas a las PCE y, si procede, una conclusión sobre la calidad y el valor generales del programa o la política correspondiente.

A fin de calibrar la calidad y el valor de sus programas y políticas, UNICEF adopta los criterios de evaluación normalizados del CAD de la OCDE³: pertinencia, eficacia, eficiencia, impacto y sostenibilidad (véase la Síntesis n.º 3 (Criterios de evaluación)). Las evaluaciones de UNICEF también deben determinar la medida en que la equidad, la igualdad de género y los derechos humanos⁴ se plasman en el diseño de la intervención, su ejecución y resultados (en el caso de las intervenciones humanitarias, quizá resulte adecuado sumar otros criterios de evaluación)⁵. Por consiguiente, las evaluaciones han de dar cabida a un

³ Organización de Cooperación y Desarrollo Económicos, Comité de Asistencia para el Desarrollo, «DAC Criteria for Evaluating Development Assistance», página web, OCDE, <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>.

⁴ Universal Management Group, *Evaluación mundial de la aplicación del enfoque basado en los derechos humanos a la programación de UNICEF*, informe final - volumen I, UNICEF, 2012.

⁵ Fondo de las Naciones Unidas para la Infancia, «Linking evaluation criteria to evaluation questions: Additional criteria for humanitarian assistance programmes», disponible en *UNICEF Monitoring and Evaluation Training Resource*, UNICEF. (En el momento en que se redacta este texto, este recurso todavía se está revisando).

conjunto de preguntas evaluativas de nivel inferior que aborden esas dimensiones específicas. Además, deben tenerse en cuenta los criterios de evaluación al plantearse cómo se define y utiliza el concepto de «éxito» en el proceso de razonamiento evaluativo.

La evaluación, por definición, debe dar respuesta a preguntas de carácter genuinamente evaluativo: no ha de interesarse únicamente por cuáles fueron los resultados (una pregunta descriptiva), sino también *por su grado de idoneidad*. Ese propósito solo es viable si se aplica el razonamiento evaluativo para valorar las pruebas relativas a las definiciones de calidad y valor.

El razonamiento evaluativo de micronivel es preciso en *todas* las evaluaciones. Consiste en contrastar las definiciones de calidad y valor con las pruebas, en busca de conclusiones sobre la idoneidad del desempeño en las dimensiones clave (por ejemplo, los resultados).

Sin el razonamiento evaluativo de micronivel, los informes de evaluación se limitan a resumir las pruebas, sin extraer conclusiones evaluativas ni responder directamente a las preguntas de evaluación. Presentar los resultados en relación con los objetivos o los indicadores establecidos tampoco supone extraer conclusiones evaluativas explícitas. Incluso cuando se emplea la evaluación basada en metas, uno de los elementos de la tarea principal consiste en determinar si el criterio de «objetivo cumplido» es realmente adecuado para concluir que se ha alcanzado un «resultado positivo». Es más, tanto si el resultado no alcanza el objetivo como si la supera ampliamente, es necesario ser capaz de determinar si este es «completamente inaceptable», «apenas aceptable», «muy bueno» o «excelente», y cómo se definen esas categorías. No basta con señalar únicamente la situación del desempeño con relación a los objetivos.

En determinados casos, se extraen conclusiones evaluativas (por ejemplo, se afirma que algo es un «resultado positivo») sin explicitar el razonamiento evaluativo correspondiente. Cuando esto ocurre, se transmite la impresión de que la evaluación se basa en una opinión personal, en lugar de en un razonamiento evaluativo sólido. Así pues, es preciso manifestar con claridad el razonamiento evaluativo, a fin de demostrar que las conclusiones están justificadas (más adelante se ofrecen ejemplos de cómo hacerlo). No solo se trata de una buena práctica profesional; un razonamiento evaluativo claro y transparente permite que otras personas entiendan la lógica y señalen posibles errores, lo que a su vez redundará en beneficio de la calidad del trabajo.

El razonamiento evaluativo de mesonivel es necesario para agrupar las pruebas referentes a múltiples dimensiones (por ejemplo, todos los resultados de una intervención o todos los aspectos relacionados con el diseño) con miras a responder a las PCE (tales como «¿en qué medida son buenos los resultados?» o «¿hasta qué punto se diseñó correctamente la intervención?»).

Muchas evaluaciones, si no la mayoría, tienden a dejar las pruebas desglosadas. Es una práctica problemática por diversas razones.

En primer lugar, *todas* las intervenciones cuentan con una combinación de puntos fuertes y débiles, historias de éxito y fracaso. Cuando un director se plantea cuál será el paso siguiente, es fundamental que sea consciente de cuáles son las deficiencias más graves, aquellas que han de resolverse con prontitud. En otras palabras, el componente de «importancia» del razonamiento evaluativo es esencial para que la evaluación resulte práctica.

En segundo lugar, aun cuando una evaluación se lleva a cabo con fines formativos (es decir, para mejorar o reconducir un programa o una política), es casi siempre importante saber, por ejemplo, si los resultados (en su conjunto, no individualmente) son lo bastante sólidos para la etapa de desarrollo correspondiente del programa o la política. Normalmente hay como mínimo una parte interesada (por ejemplo, un financiador) que necesita saber lo antes posible si la inversión se encuentra en el buen camino con miras a la consecución de resultados valiosos.

En tercer lugar, se observan con frecuencia «respuestas» a las PCE que no se generaron mediante una síntesis del desempeño de la intervención respecto a todos los criterios, sino a partir de la elección de uno o dos indicadores que se consideraron los más importantes y se presentaron como «la respuesta».

Aunque es válido otorgar *un peso mayor* (cuantitativa o cualitativamente) a los aspectos más importantes, no lo es escoger un pequeño número de indicadores *en detrimento* de otras dimensiones de mérito.

El razonamiento evaluativo de macronivel es necesario siempre que se plantea una pregunta global sobre el valor de un programa o una política en su totalidad. Esa pregunta global se responde al sintetizar las respuestas a todas las PCE, incluidas aquellas sobre el diseño y la ejecución de la intervención, sus resultados previstos e imprevistos, y su rentabilidad comparada.

En resumen, en las [evaluaciones de impacto](#):

- el razonamiento evaluativo es *siempre* necesario en el micronivel, a fin de determinar el grado de idoneidad de los resultados en dimensiones concretas
- el razonamiento evaluativo es siempre necesario en el mesonivel, a fin de responder adecuadamente a las PCE establecidas de antemano en la evaluación
- el razonamiento suele ser necesario (pero no lo es en todos los casos) en el macronivel, a fin de llegar a una conclusión general sobre si la intervención es una inversión provechosa de tiempo, esfuerzo, dinero y otros recursos.

3. ¿CÓMO LLEVAR A CABO UN RAZONAMIENTO EVALUATIVO?

Llevar a cabo un razonamiento evaluativo resulta más sencillo cuando se integra desde el primer momento en el diseño de las evaluaciones. Aunque es posible hacerlo retrospectivamente, resulta más complicado. Si cambian las necesidades o prioridades, o si surgen resultados nuevos o imprevistos, es preciso efectuar el razonamiento evaluativo en «tiempo real».

Los pasos que se describen a continuación rastrean el razonamiento evaluativo desde la fase de diseño de la evaluación, aunque se indican las opciones disponibles cuando el proceso se incorpora en una fase posterior de la evaluación y debe retroadaptarse a los datos ya recopilados. Todos los pasos siguientes pueden darse de manera colaborativa (con la participación de los interesados) o independiente (por parte del equipo de evaluación).

1. **Desarrolle un conjunto de PCE en los que se fundamentará la evaluación en su totalidad.** Esta tarea debería acometerse al determinar el alcance de la evaluación y planificarla (a ser posible, a partir de una versión provisional de los términos de referencia). Según la experiencia de la autora, 7 ± 2 preguntas suele ser un número apropiado. Permite abarcar distintos aspectos del programa o política, pero es lo suficientemente bajo como para no abrumar al lector. Cada una de esas preguntas de alto nivel se desglosará en un grupo de subpreguntas.
2. **Identifique el criterio adecuado para definir un «resultado valioso» y un «buen programa o política» en este contexto.** En concreto, tenga en cuenta las dimensiones relacionadas con los criterios de evaluación mencionados anteriormente (por ejemplo, los criterios del CAD de la OCDE o el enfoque de programación basado en los derechos humanos). Las fuentes principales serán en este caso las evaluaciones de puntos fuertes y débiles, los valores y las ambiciones de la comunidad, y los estándares pertinentes al tipo de iniciativa, entre otras.

También resultarán útiles los documentos de planificación del programa o política, en los que se reflejan sus objetivos y estrategias fundamentales. No obstante, para determinar si un resultado es «bueno», será preciso establecer un criterio más sólido que el de «objetivo cumplido» —y, cuando no se consigue, también es necesario poder indicar, por ejemplo, en qué medida se ha estado cerca de alcanzar el objetivo—.

3. **Conciba una o varias rúbricas evaluativas que definan la combinación de pruebas en distintos planos del desempeño.** Una rúbrica evaluativa es una tabla en la que se describen las

pruebas en distintos planos del desempeño, a partir de un criterio de interés determinado o con respecto al programa o política en su conjunto.

La clave reside en no limitarse a crear grupos sobre un indicador cuantitativo o una puntuación (por ejemplo, entre el 90% y el 100% = excelente), sino debatir qué *combinación de pruebas* (cualitativas, cuantitativas o ambas) deberían observarse en cada plano del desempeño.

En ocasiones será necesario crear una rúbrica basada en un único indicador muy importante. Incluso en esos casos, es fundamental plantearse qué criterios se aplicarán a las pruebas no solo en términos generales, sino también para ciertos subgrupos (por ejemplo, suele ser importante examinar de forma especial los resultados en favor de los grupos de población vulnerables), a fin de determinar si se ha obtenido un «buen» resultado (es decir, provechoso, equitativo, muy necesario, etc.). En la mayoría de los casos, la inclusión de pruebas cualitativas permitirá «dar vida» a los datos cuantitativos y hacerse una idea más clara de lo que el resultado supone en la práctica.

4. **Diseñe instrumentos y protocolos de recopilación de datos con los que reunir las pruebas más determinantes para la definición de calidad, según se establece en la rúbrica evaluativa.** La rúbrica traza la imagen general de la combinación de pruebas cualitativas y cuantitativas, y de ese modo permite tener una idea más clara de todo lo necesario para determinar cómo se calificará el desempeño.

En aquellas evaluaciones en que deban emplearse los datos existentes o ya se hayan reunido las pruebas, es importante no centrarse exclusivamente en el material disponible al redactar la rúbrica; en su lugar, debemos trazar una imagen amplia del desempeño, independientemente de las pruebas a nuestra disposición.
5. **Lleve a cabo un razonamiento evaluativo de micronivel.** Interprete las pruebas en relación con las definiciones de calidad y valor, a partir de la [rúbrica evaluativa](#) u otra directriz adecuada de interpretación de la evaluación.
6. **Recurra a la evaluación de necesidades u otra prueba pertinente para determinar qué aspectos son de importancia esencial; cuáles son importantes, pero no fundamentales; y cuáles resultan beneficiosos pero no son tan significativos.** Al igual que los demás pasos del proceso, puede efectuarse en colaboración con los interesados o de forma independiente por parte del equipo de evaluación.
7. **Partiendo de la evaluación de importancia anterior, establezca cómo se sintetizarán las pruebas relativas a múltiples dimensiones a fin de generar respuestas a las PCE.** No se trata simplemente de asignar puntuaciones y pesos (como en el enfoque de ponderación y suma numérica). El razonamiento evaluativo de mesonivel debe fundamentarse en criterios, de manera que sea posible incorporar «umbrales». Los umbrales son niveles mínimos de desempeño en determinadas dimensiones clave, indispensables para que el programa o política se consideren mínimamente aceptables en relación con una pregunta de evaluación (un «umbral firme») o alcancen o superen una calificación concreta (un «umbral flexible»)⁶.
8. **Basándose en sus conocimientos especializados pertinentes, el equipo de evaluación debe considerar las pruebas y llevar a cabo un razonamiento evaluativo de mesonivel.** El producto final de ese proceso debe consistir en una serie de respuestas evaluativas directas y explícitas a cada una de las PCE que constituyen el marco general de la evaluación. Esas respuestas directas han de sintetizarse en el resumen del informe.
9. **Si también fuera necesario efectuar un razonamiento evaluativo de macronivel, genere y aplique una rúbrica general para el programa o política que describa cuál debe ser su desempeño global, teniendo en cuenta las respuestas a todas las PCE.** De nuevo, este proceso

⁶ Davidson, E. Jane, *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*, Sage, Thousand Oaks, 2005.

no debe consistir en un ejercicio de ponderación y suma numérica, sino en el esbozo de los criterios sobre el conjunto de desempeños a partir de los que se determinará que un programa o política fueron enormemente exitosos y valiosos (y no simplemente adecuados ni «fallidos»).

En la medida de lo posible, es muy recomendable que las rúbricas correspondientes a los tres niveles de razonamiento evaluativo se elaboren *antes* de reunir las pruebas. De ese modo, será posible reflexionar y debatir con más claridad sobre la definición de éxito, sin caer en la tentación de arrojar una luz positiva sobre las pruebas disponibles.

La participación de los interesados en la elaboración de las rúbricas evaluativas no solo mejora su validez y credibilidad, sino que además permite a los responsables de las decisiones imaginar cómo serán las pruebas y comprender por anticipado cómo se interpretarán. Se reducen así los malentendidos y desacuerdos en la interpretación posterior de las pruebas y, por consiguiente, se acepta en mayor grado que las constataciones son válidas y creíbles, y aumenta el compromiso con las medidas que sea necesario adoptar.

Las rúbricas evaluativas deben tratarse siempre como instrumentos vivos y aplicarse a **los nuevos resultados y a las prioridades cambiantes**, así como a aquellos que se prevén desde el primer momento.

Cuando las necesidades o prioridades cambien, las rúbricas deben revisarse para reflejar este hecho. En ocasiones, habrá que recopilar pruebas diferentes que permitan entender mejor el grado de desempeño del programa o política.

Cuando se produzcan resultados nuevos o imprevistos, seguirá siendo necesario determinar la medida en que resultaron beneficiosos o perjudiciales con respecto a las necesidades y ambiciones de la comunidad, al igual que sucede con los resultados deseados y previstos. En esos casos, deben elaborarse rúbricas en cuanto se identifiquen los nuevos resultados, aunque es deseable hacerlo antes de reunir las pruebas para que sea posible racionalizar en tiempo real la recopilación de pruebas y disponer de un marco viable de interpretación cuando estas se obtienen.

4. ASPECTOS PRÁCTICOS

Los principales aspectos que hay que tener en cuenta al planificar un razonamiento evaluativo son:

- ¿Cómo pueden definirse de forma adecuada la calidad y el valor en este contexto, para este programa o política? ¿Por qué? ¿Qué opiniones habría que escuchar al respecto?
- ¿Se han diseñado ya los instrumentos de recopilación de datos para llevar a cabo el razonamiento evaluativo con el material disponible? ¿O es demasiado pronto para que el razonamiento evaluativo oriente el diseño de los instrumentos de recopilación de datos?
- ¿Quién debería participar en el propio proceso de razonamiento evaluativo? Aspectos principales:
 - **Validez:** ¿Quién dispone de los conocimientos especializados necesarios para llevar a cabo correctamente el razonamiento evaluativo? Además de un evaluador principal con experiencia en la metodología específica de la evaluación, será necesario incorporar a personas con conocimientos sobre el asunto que se trata; sobre el contexto y la cultura locales; y sobre la interpretación de los distintos tipos de pruebas que se emplearán.
 - **Credibilidad:** ¿Quién debe participar en el razonamiento evaluativo a fin de que sus constataciones resulten creíbles a otras personas?
 - **Utilidad:** ¿Quién empleará con mayor probabilidad las constataciones de la evaluación (es decir, los productos de la síntesis evaluativa)? Puede resultar útil contar en el análisis con la participación de los usuarios principales previstos, que de ese modo podrán observar los

resultados con sus propios ojos. Aumentan así la comprensión de las constataciones y el compromiso con su aplicación.

- **Voz:** ¿Quién tiene derecho a estar presente cuando se lleve a cabo el razonamiento evaluativo? Es un aspecto particularmente importante cuando se trabaja con grupos de personas a los que se ha excluido tradicionalmente de la mesa de evaluación, tales como los pueblos indígenas. ¿Deberían participar los destinatarios del programa o política, o representantes de la comunidad? ¿Y los financiadores? ¿Y los directivos de los programas?
- **Costo:** Sin olvidar el costo de oportunidad que acarrea que el personal abandone su trabajo para participar en la evaluación, ¿en qué etapas del proceso de razonamiento evaluativo es preferible contar con distintos interesados, con miras a lograr el mayor beneficio mutuo posible?

Una dificultad notable reside en que el razonamiento evaluativo, si bien esencial, no es todavía un conjunto de aptitudes habitual en el repertorio de los profesionales de la evaluación, pues no figura en el plan de estudios de prácticamente ningún programa de capacitación avanzada o posgrado en evaluación. No resulta sorprendente, por tanto, que las personas que encargan y supervisan el trabajo de evaluación aún no comprendan plenamente esta cuestión.

Un segundo reto significativo lo encontramos al definir qué constituye una intervención «buena», que entre otros aspectos supone plantearse si la propia intervención es ética. Se trata de otro ámbito en el que la capacitación sobre evaluación es profundamente deficiente, por lo que la mayoría de los equipos de evaluación están poco preparados para llevar a cabo un análisis ético de los programas o las políticas.

Es probable que los directivos que deseen contratar a un equipo de evaluación con este conjunto de aptitudes requieran ayuda para establecer con claridad en los términos de referencia sus expectativas referentes al razonamiento evaluativo, entre otras la necesidad de un análisis ético de la propia intervención. Los directivos quizá también precisen asesoramiento experto y apoyo a lo largo de los procesos de contratación y gestión de la evaluación para negociar y administrar con eficacia todo lo necesario, incluso con contratistas que todavía estén haciéndose con la metodología.

5. ¿QUÉ OTROS MÉTODOS FUNCIONAN BIEN CON ESTE?

El razonamiento evaluativo es obligatorio en todas las evaluaciones, con independencia de los métodos o el enfoque aplicados. Lo más importante es que *todas* las evaluaciones tienen la misión de plantear y responder a unas preguntas de evaluación —no deben limitarse a medir y describir resultados—. Así pues, todas las conclusiones han de interpretarse en un marco evaluativo que no solo indique cuáles son los resultados, sino también su grado de idoneidad. Es necesario que el razonamiento evaluativo sea capaz de hacerlo.

6. PRESENTACIÓN DE LOS RESULTADOS Y ANÁLISIS

La estructura de un informe de evaluación puede alentar en gran medida la presentación resumida de respuestas directas a las preguntas de evaluación, respaldadas por un grado de detalle suficiente sobre el razonamiento evaluativo y la metodología, de manera que el lector comprenda la lógica y observe con claridad la base empírica.

Las recomendaciones siguientes ayudarán a establecer expectativas claras para los informes de evaluación fuertemente apoyados en el razonamiento evaluativo:

1. El resumen debe ofrecer respuestas evaluativas directas y explícitas a las PCE en las que se fundamenta el conjunto de la evaluación.

2. Ha de emplearse un lenguaje evaluativo explícito al exponer las constataciones (no un lenguaje neutro que se limite a describirlas). Deben facilitarse ejemplos (por ejemplo, del tipo que se observa en el ejemplo de buenas prácticas a continuación).
3. Deben usarse sistemas de visualización de datos claros y sencillos con «instantáneas» de fácil comprensión sobre el desempeño de la intervención en distintas dimensiones de mérito.
4. Ha de estructurarse la sección de constataciones mediante el empleo de las PCE como subepígrafes (en lugar de los tipos y fuentes de pruebas que se utilizan con frecuencia).
5. El razonamiento evaluativo aplicado ha de describirse de forma clara y transparente, mediante explicaciones que puedan entender con facilidad tanto las personas que no son evaluadores como los lectores que carezcan de conocimientos especializados en la materia. Esas explicaciones deben ser de amplio alcance y breves en el cuerpo principal del informe, y más pormenorizadas en los anexos.
6. Si la extensión de las rúbricas evaluativas es relativamente reducida, deben incluirse en el cuerpo principal del informe. De lo contrario, deben resumirse al menos una o dos en el cuerpo principal e incluir todas las rúbricas completas en un anexo.

Uno de los sellos distintivos de un gran razonamiento evaluativo es la capacidad de transmitir de modo sucinto y claro los aspectos importantes sin necesidad de pasar por alto detalles significativos.

7. EJEMPLOS DE BUENAS PRÁCTICAS

Dado que pocas evaluaciones de impacto en el desarrollo emplean de forma sistemática y expresa el razonamiento evaluativo para responder a las preguntas de evaluación, el ejemplo del cuadro 2 se refiere en su lugar a una evaluación educativa de un programa de liderazgo en la escuela.

Este estudio de caso ilustra el uso de las PCE y las rúbricas evaluativas con miras a la interpretación de pruebas. Cada pregunta de alto nivel es expresamente evaluativa, pues no se limita al «qué» y al «cómo», sino que también se plantea «con qué grado de idoneidad».

Cuadro 2. Estudio de un caso de buenas prácticas: Evaluación del programa de orientación para directores novatos

Esta evaluación se estructuró en torno a cuatro PCE; la principal era:

¿En qué medida es el programa de orientación para directores novatos un programa de gran calidad que orienta a los directores novatos de los centros de primaria, secundaria y las escuelas kura kaupapa (centros de inmersión en la lengua indígena maorí) a fin de que asuman el liderazgo educativo en sus escuelas?

Las rúbricas se elaboraron de manera colaborativa. Se hizo un hincapié especial en la presencia de la perspectiva maorí (indígena) a lo largo del proceso de razonamiento evaluativo, importante de cara a su validez, credibilidad, justicia y posterior utilización. A continuación se ofrece una muestra de los dos niveles superiores de una de las rúbricas. Cabe destacar la incorporación en la rúbrica de una serie de aspectos del desempeño, así como una declaración general inequívoca en la parte superior de cada categoría de calificación que explica su propósito global.

Rúbrica sobre la participación en el programa de capacitación profesional y reconocimiento como «líderes pedagógicos» en sus centros

Calificación	Criterios de evaluación
Excelente	<p>Ejemplo inequívoco de desempeño ejemplar o mejor práctica en su ámbito; carece de puntos débiles</p> <ul style="list-style-type: none"> • Se establecen siempre de forma clara y adecuada los objetivos de desarrollo profesional de los directores novatos • Se establece en todo momento una buena relación de trabajo entre los directores novatos y los tutores, que les brindan asesoramiento y consejo profesional • Se crean siempre redes de apoyo entre pares entre los directores novatos y sus homólogos • Los directores novatos se implican en un aprendizaje reflexivo sobre la asunción del liderazgo pedagógico en sus centros • Los directores novatos entienden siempre la importancia de asumir el liderazgo pedagógico y contar con estrategias claras a tal efecto • Los directores novatos señalan en todo momento que saben cómo recopilar, analizar y aplicar los datos para respaldar el aprendizaje de sus alumnos • Se dispone siempre de pruebas acerca de la atención que prestan los directores novatos a la equidad para los maoríes • El respaldo a los directores novatos se coordina adecuadamente (especialmente cuando trabajan a la vez con varios grupos de apoyo).

Muy buena	<p>Desempeño muy bueno o excelente en prácticamente todos los aspectos; sólido en general, aunque no ejemplar; sin puntos débiles con verdadera repercusión</p> <ul style="list-style-type: none"> • Se establecen casi siempre de forma clara y adecuada los objetivos de desarrollo profesional de los directores novatos • Se establece casi siempre una buena relación de trabajo entre los directores novatos y los tutores, que les brindan asesoramiento y consejo profesional • Se crean casi siempre redes de apoyo entre pares entre los directores novatos y sus homólogos • Los directores novatos se implican casi siempre en un aprendizaje reflexivo sobre la asunción del liderazgo pedagógico en sus centros • Los directores novatos entienden casi siempre la importancia de asumir el liderazgo pedagógico y contar con estrategias claras a tal efecto • Los directores novatos señalan casi siempre que saben cómo recopilar, analizar y aplicar los datos para respaldar el aprendizaje de sus alumnos • Se dispone casi siempre de pruebas acerca de la atención que prestan los directores novatos a la equidad para los maorías • El respaldo a los directores novatos se coordina casi siempre adecuadamente (especialmente cuando trabajan a la vez con varios grupos de apoyo).
-----------	--

Fuente: Reproducido bajo una licencia Creative Commons de Oakden, Judy, *Evaluation rubrics: how to ensure transparent and clear assessment that respects diverse lines of evidence*, BetterEvaluation, 2013, p. 12. Véase <http://betterevaluation.org/resource/example/rubrics-oakden>.

Las rúbricas de calidad como esta fomentan el uso de criterios sólidos de evaluación que facilitan un lenguaje común para interpretar las pruebas e incrementan su [fiabilidad](#) y [validez](#) entre evaluadores. Este planteamiento contrasta con numerosos enfoques de «cuantificación» que intentan eliminar el criterio personal como si fuera intrínsecamente poco fiable y se fundamentase exclusivamente en la opinión.

La evaluación también recurrió a otras buenas prácticas, tales como un proceso de fases múltiples para evaluar los datos en relación con las rúbricas. Todos los datos cualitativos (es decir, procedentes de las entrevistas con los interesados clave) y cuantitativos (de una encuesta entre los directores) se transformaron en calificaciones —en una escala de excelente a deficiente— referentes a cada rúbrica, como se aprecia en la tabla 2. Este proceso de síntesis permitió efectuar juicios evaluativos generales sobre los resultados basados en todas las fuentes de datos⁷.

⁷ Para obtener más información sobre el proceso, véase Oakden, Judy, *Evaluation rubrics: How to ensure transparent and clear assessment that respects diverse lines of evidence*, BetterEvaluation, 2013, p. 10. Disponible en Internet: <http://betterevaluation.org/sites/default/files/Evaluation%20rubrics.pdf>.

Tabla 2. Calificaciones empleadas para evaluar los datos cuantitativos y cualitativos en relación con las rúbricas

Calificación	Datos cuantitativos y cualitativos
Excelente: siempre	Ejemplo inequívoco de desempeño ejemplar o mejor práctica en su ámbito; carece de puntos débiles. Es probable que al menos el 90% coincida con esta afirmación en un grado notable o elevado.
Muy buena: casi siempre	Desempeño muy bueno o excelente en prácticamente todos los aspectos; sólido en general, aunque no ejemplar; sin puntos débiles con verdadera repercusión. Es probable que entre el 80% y el 90% coincida con esta afirmación en un grado notable o elevado.
Buena: en su mayor parte, con algunas salvedades	Desempeño general razonablemente bueno; quizá presente leves deficiencias, pero nada grave. Entre el 60% y el 80% coincide con esta afirmación en un grado notable o elevado, y no más del 15% coincide de forma limitada o muy limitada.
Adecuada: en ocasiones, con bastantes salvedades	Desempeño aceptable, con ciertas deficiencias graves pero no funestas en algunos aspectos. Entre el 40% y el 60% coincide con esta afirmación en un grado notable o elevado, y no más del 15% coincide de forma limitada o muy limitada.
Deficiente: nunca (u ocasionalmente, con deficiencias manifiestas)	Pruebas inequívocas de un funcionamiento insatisfactorio; deficiencias graves generalizadas en aspectos fundamentales. Probablemente menos del 40% coincida con esta afirmación en un grado notable o elevado.

Fuente: Reproducido bajo una licencia Creative Commons de Oakden, J. (2013). *Evaluation rubrics: How to ensure transparent and clear assessment that respects diverse lines of evidence*. Better Evaluation, p. 10. Disponible en Internet: <http://betterevaluation.org/sites/default/files/Evaluation%20rubrics.pdf>; Oakden adaptó el contenido de la tabla de Davidson, E. Jane, *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*, Sage, Thousand Oaks, 2005.

8. EJEMPLOS DE POSIBLES PROBLEMAS

Algunas evaluaciones no aplican adecuadamente el razonamiento evaluativo y en su lugar exponen una serie de constataciones no evaluativas de manera descriptiva, sin extraer conclusiones claras sobre el grado de idoneidad general del desempeño.

A fin de ilustrar algunos de los problemas que acarrea no aplicar adecuadamente el razonamiento evaluativo, se ofrece como ejemplo la evaluación de un programa de desarrollo juvenil financiado por UNICEF. El extracto (modificado ligeramente) del resumen muestra que se presentan una serie de hechos no evaluativos sin comentarios referentes a su mérito o valor:

En el marco del programa, 14.000 jóvenes recibieron capital inicial para poner en marcha sus propias empresas. El 89% de los que habían iniciado un negocio propio lo dirigían con éxito. El porcentaje de mujeres que dirigían con éxito su empresa era superior al de hombres. Las causas principales de un mal funcionamiento de las empresas fueron el precio elevado de los insumos y la inexistencia de un mercado para sus productos. El 76% de los jóvenes que solicitaron un crédito para poner en marcha su negocio señaló que su vida había mejorado gracias a la creación de una pequeña empresa. En lo que respecta a la provisión de capital inicial, el costo por persona del apoyo económico osciló entre

45 y 250 dólares. En la mayoría de las regiones, el costo por persona apenas aumentó, o no lo hizo en absoluto. Este aspecto puede relacionarse con el incremento del costo de los insumos en el país a lo largo del periodo.

Tanto en términos de capacitación como de provisión de capital, el número de hombres beneficiarios duplicó ampliamente el de mujeres.

El resumen presenta hechos y datos sobre el número de participantes, compara los índices de participación y los resultados positivos de hombres y mujeres, el porcentaje de participantes que tuvo éxito y el costo por persona, pero no establece la calidad ni el valor de tales resultados.

Para el lector interesado en las conclusiones evaluativas del informe, sigue sin darse respuesta a las siguientes preguntas:

- ¿Es la cifra de 14.000 jóvenes destinatarios de capital inicial adecuada, dado el volumen de la inversión en este componente del programa?
- ¿Y la composición de género —con más del doble de hombres que de mujeres beneficiarios—, especialmente ante el hecho de que las mujeres mostraron más probabilidades de éxito empresarial?
- ¿En qué medida es positivo que el 89% de los participantes dirigiera su propia empresa en el momento específico analizado?
- ¿Obtuvo este componente del programa resultados rentables, dado el costo por persona del apoyo económico?

Cuando se exponen resultados de forma no evaluativa, como en el ejemplo, la calidad o el valor del programa o política no se aprecian con claridad en el resumen. Al contrario, el lector se ve obligado a aplicar su propio criterio a las pruebas para dar sentido a las constataciones.

La inexistencia de un razonamiento evaluativo también hace difícil apreciar la relación lógica entre las cifras y descripciones, por una parte, y el listado de recomendaciones que se facilita al principio del informe.

9. LECTURAS Y ENLACES CLAVE

BetterEvaluation, «Synthesise data from a single evaluation», página web, BetterEvaluation, http://betterevaluation.org/plan/synthesize_value/synthesize_data_single_evaluation. (Incluye un documento de dos páginas descargable en PDF en el que se resumen las principales opciones de cara a para llevar a cabo una síntesis evaluativa).

The Center for Advanced Research on Language Acquisition, «Process: Creating rubrics», página web, CARLA, 2012, <http://www.icyte.com/saved/www.carla.umn.edu/560426>.

Coyle, J., *et al.*, «Evaluation rubrics with E. Jane Davidson», serie de podcasts *Adventures in Evaluation*, 2013, <http://adventuresinevaluation.podbean.com/2013/12/08/evaluation-rubrics-with-e-jane-davidson/>.

Dart, Jessica, *et al.*, *Review of Evaluation in Agricultural Extension*, Rural Industries Research and Development Corporation, publicación de la RIRDC n.º 98/136, 1998.

Davidson, E. Jane, *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*, Sage, Thousand Oaks, 2005.

Davidson, E. Jane, *Principios básicos de la evaluación para la acción: Obteniendo respuestas sucintas a las preguntas más importantes*, minilibro, Real Evaluation, 2013. Disponible como libro de bolsillo, PDF o libro electrónico en inglés, español y francés. Véase <http://realevaluation.com/read/minibooks/>.

Evaluate Research, *Final Evaluation Report of the Recognised Seasonal Employer Policy (2007–2009)*, Ministerio de Trabajo, Gobierno de Nueva Zelandia, Wellington, 2010. Véase http://www.immigration.govt.nz/NR/rdonlyres/D7CDDA62-DCF6-4219-AC69-29B88F9C5513/0/web_DOL11246RSEEvaluation_v3.pdf.

Oakden, Judy, «Evaluation rubrics: how to ensure transparent and clear assessment that respects diverse lines of evidence», página web, BetterEvaluation, 2013, <http://betterevaluation.org/resource/example/rubrics-oakden>.

Parcell, Geoff y Chris Collison, *No More Consultants: We know more than we think*, John Wiley & Sons, Londres, 2009. (Un útil diagrama de río se resume en http://betterevaluation.org/resources/tools/rubrics/river_chart).

Rogers, Patricia J. y E. Jane Davidson, blog con diversos artículos sobre las rúbricas evaluativas, Genuine Evaluation, <http://genuineevaluation.com/?s=rubrics>.

Scriven, Michael, *Evaluation Thesaurus*, cuarta edición, Sage, Newbury Park, 1991. (Fuente primordial para entender la lógica subyacente en las síntesis evaluativas).

GLOSARIO

<u>Crterios de evaluacón</u>	<i>Valores (principios, atributos o cualidades considerados intrínsecamente buenos, deseables, importantes y de valor general) que se utilizarán en una evaluacón para juzgar el mérito de una intervencón. Algunos ejemplos son los criterios del CAD de la OCDE, el enfoque de programacón basado en los derechos humanos y los criterios sobre la asistencia humanitaria.</i>
Datos cualitativos	<i>Datos descriptivos que es posible observar, pero no medir. Entre otros, textos, imágenes y sonidos, pero no valores numéricos o cuantitativos. Véase: datos cuantitativos.</i>
Datos cuantitativos	<i>Medicón de valores o recuentos expresados en forma numérica. Los datos cuantitativos pueden cuantificarse, verificarse y emplearse en el análisis estadístico. Los resultados suelen ser generalizables, aunque no en todos los casos. Véase: datos cualitativos.</i>
<u>Evaluacón de impacto</u>	<i>Evaluacón que proporciona información sobre los impactos que produce una intervencón. Puede realizarse una evaluacón de impacto de un programa o una política o del trabajo preliminar, como la creacón de capacidad, la promoción de políticas y el apoyo a la creacón de un entorno propicio. Supone examinar no solo las metas y los objetivos, sino también los impactos imprevistos. Véase: impacto.</i>
<u>Evaluacón de procesos</u>	<i>Evaluacón de la dinámica interna de los organismos de ejecución, sus instrumentos de política, sus mecanismos de prestación de servicios, sus procedimientos de gestión y los vínculos que existen entre todos estos componentes. (Definición del CAD de la OCDE, 2010) Véase: evaluacón formativa.</i>
<u>Evaluacón evolutiva</u>	<i>Enfoque de evaluacón dirigido a comprender las actividades de un programa que se desarrolla en un entorno complejo y dinámico, caracterizado por la flexibilidad metodológica y el pensamiento sistémico. Se centra en la innovación y el aprendizaje estratégico, en lugar de en los resultados estándar. Facilita el intercambio de impresiones en tiempo real con los interesados en el programa y fomenta el desarrollo, el reajuste y la adaptación continuos.</i>
<u>Evaluacón formativa</u>	<i>Evaluacón cuyo objeto es mejorar el desempeño y que generalmente se realiza durante la fase de implementación de un proyecto o programa. También pueden realizarse evaluaciones formativas por otras razones como la verificación del cumplimiento, por exigencias legales o como parte de una iniciativa de evaluacón de mayor envergadura. Véase: evaluacón de procesos.</i>
<u>Evaluacón sumativa</u>	<i>Estudio que se realiza al final de una intervencón (o de una fase de esa intervencón) para determinar en qué medida se produjeron los resultados previstos. El objetivo de la evaluacón sumativa es proporcionar información sobre el valor del programa.</i>
<u>Fiabilidad (entre evaluadores)</u>	<i>Medicón de la fiabilidad que evalúa el grado de acuerdo entre las decisiones o calificaciones de distintos evaluadores, jueces o codificadores. Establece una puntuación sobre el consenso en las calificaciones de los evaluadores. Términos conexos: acuerdo entre evaluadores.</i>
<u>Impacto</u>	<i>Efectos de largo plazo positivos y negativos, primarios y secundarios, producidos directa o indirectamente por la intervencón, intencionalmente o no. (Definición del CAD de la OCDE, 2010)</i>

<u>Investigación cualitativa</u>	<i>Investigación que recopila y analiza datos cualitativos. Entre sus métodos tradicionales figuran los estudios de casos, la observación y la etnografía. Sus resultados no suelen considerarse generalizables, pero sí son a menudo transferibles. Véase: datos cualitativos, datos cuantitativos, investigación cuantitativa.</i>
<u>Investigación cuantitativa</u>	<i>Investigación que recopila y analiza datos cuantitativos. Véase: datos cuantitativos, datos cualitativos, investigación cualitativa.</i>
<u>Investigación primaria</u>	<i>Investigación en la que se recopilan y analizan datos nuevos a efectos de un proyecto determinado. Cabe mencionar como ejemplos las encuestas, las entrevistas y las observaciones. Véase: investigación secundaria.</i>
<u>Investigación secundaria</u>	<i>Análisis de los datos disponibles que no se recopilan principalmente a los efectos del proyecto de investigación o la evaluación correspondiente. Es posible que los datos se reúnan en el marco de otro proyecto de investigación o evaluación, la elaboración de un censo, el seguimiento del desempeño u otro fin. Véase: investigación primaria</i>
<u>Métodos mixtos</u>	<i>Métodos de investigación que recopilan o analizan datos tanto cuantitativos como cualitativos para estudiar una misma cuestión. Los métodos mixtos permiten triangular datos y adoptar varias perspectivas de investigación. Véase: investigación cualitativa, investigación cuantitativa.</i>
<u>Preguntas clave de evaluación</u>	<i>Preguntas de evaluación de alto nivel (macronivel) sobre el desempeño general, a las que la evaluación debe tratar de dar respuesta. Las preguntas clave de evaluación se derivan del propósito de la evaluación.</i>
<u>Preguntas evaluativas</u>	<i>Las preguntas evaluativas se refieren a la conclusión general sobre si un programa o una política puede considerarse un éxito, una mejora o la mejor opción.</i>
<u>Resultado</u>	<i>Efectos a corto y mediano plazo probables o logrados de los productos de un programa o política; por ejemplo, un cambio en los niveles de vacunación o los comportamientos clave.</i>
<u>Rúbrica evaluativa</u>	<i>Tabla en la que se describen las pruebas en distintos planos del desempeño, a partir de un criterio de interés determinado o con respecto a la intervención (programa o política) en su conjunto.</i>
<u>Términos de referencia</u>	<i>Descripción de los antecedentes, los objetivos, los usuarios previstos, las preguntas clave de evaluación, la metodología, las funciones y responsabilidades, los plazos, los entregables, las normas de calidad, las aptitudes del equipo de evaluación y otras cuestiones pertinentes en las que se fundamenta un contrato de UNICEF con los evaluadores.</i>
<u>Validez</u>	<i>Medida en que un estudio refleja o valora con precisión el concepto específico que el investigador pretende medir. Cabe la posibilidad de que un método sea fiable, al medir de manera coherente el mismo concepto, pero no válido. Véase: fiabilidad.</i>