

3ª Conferencia Internacional sobre las Capacidades Nacionales de Evaluación: Soluciones innovadoras a los desafíos relacionados con la independencia, la credibilidad y el uso de las evaluaciones

Documento Temático sobre la Credibilidad¹

María Bustelo
Sociedad Europea de Evaluación
Universidad Complutense de Madrid
mbustelo@cps.ucm.es

En este trabajo se pretende abrir un debate sobre la credibilidad de las evaluaciones y sobre la manera de mejorar y garantizar su credibilidad. Muchos problemas afectan a la credibilidad de la evaluación -la especialización e independencia de los evaluadores, el grado de transparencia en el proceso de evaluación, y la calidad de los productos, por nombrar sólo algunos. El contexto cultural también es importante - los valores sobre los que descansa la evaluación, la forma en que la evidencia se considera creíble, las instituciones que apoyan los sistemas y estructuras de evaluación, las personas que contribuyen a la evaluación, y cómo la evaluación es compartida, comunicada y presentado su informe.

Mi argumento al escribir este documento proviene de mi experiencia en la formación de evaluadores y en la dirección del Máster en Evaluación de Programas y Políticas Públicas de la Universidad Complutense de Madrid (España), que comenzará su 12º curso en el ciclo 2013-2014. También viene de mi servicio en el Consejo Directivo de la Sociedad Europea de Evaluación (EES) desde 2009, que en los dos últimos años he tenido el honor de dirigir desde la presidencia de esta sociedad regional. Para brindar contextualizar este documento, utilizaré el caso de España, que es mi propio país y donde he estado trabajando en evaluación los últimos 25 años, así como también a la región europea más amplia y diversa. Y me voy a referir más de una vez a la inspiradora Audiencia Pública del EES en el Parlamento Europeo sobre "La evaluación en la Democracia", celebrada en Bruselas en abril de 2013.²

1. La naturaleza política de la evaluación y la necesidad de credibilidad

La naturaleza política de la evaluación ha sido reconocida ampliamente por la comunidad de evaluación desde finales de los 1980s. En su excelente discurso en la reunión anual de la Asociación Americana de Evaluación en San Diego en 1997, que fue publicado posteriormente como un artículo (Chelimsky, 1998), Eleanor Chelimsky habló sobre "el papel de la experiencia en la formulación de las teorías de la práctica de la evaluación". Ella afirma que la evaluación y la política están "visceralmente conectadas". Para ella, "las evaluaciones han sido profundamente

¹ El presente documento es la traducción del trabajo "Guiding Thematic Paper on Credibility". La traducción fue realizada gracias al trabajo voluntario de Brenda Lora, Eduardo Centeno, Carlos Frías Coronado, y Pablo Rodríguez-Bilella, miembros de la ReLAC –Red de Seguimiento, Evaluación y Sistematización de América Latina y el Caribe- <http://noticiasrelac.ning.com/>

² www.czech-in.org/ees/ees-newsletter-2013-06-june-special.pdf

afectados por la forma en que funciona la política en una sociedad democrática y pluralista, y sobre todo por la naturaleza continuamente cambiante de la situación política". Además, debido a que "en un mundo de maniobras altamente sofisticadas y continuas para obtener ventajas políticas, el abogar (*advocacy*) abunda. No sólo los políticos tienen sus propias agendas políticas, también están asediados por los grupos de presión, los intereses creados y grupos de lobby, todos con sus historias de guerra sobre el "éxito" o "fracaso", y todos tratando, con el dinero, el poder, y los datos, mover las políticas y programas en direcciones específicas "(Chelimsky, 1998: 37-38). Por lo que necesitamos tener en cuenta e incluir esta naturaleza política en el corazón mismo del concepto y las teorías de la evaluación. Esa necesidad de la credibilidad es particularmente importante al considerar la política como elemento central de las prácticas de evaluación. Para Chelimsky, la credibilidad implica imparcialidad y eso debe ser algo que se conserva en el proceso de evaluación, mientras que las evaluaciones deben ser percibidas como imparciales. Así que no sólo deberían ser imparciales, deben ser vistas como imparciales por los interesados. En su amplia experiencia en la Oficina de Contabilidad General de EE.UU., esta imparcialidad era a menudo preservada corrigiendo posibles percepciones de parcialidad de un evaluador, mediante la contratación de la tendencia opuesta en el mismo equipo de evaluación, así como mediante la corrección de este sesgo potencial durante el proceso de evaluación, especialmente en la revisión de la literatura, el diseño metodológico y el informe de las fases de evaluación. Estos procedimientos fueron útiles en un contexto como el de la Oficina de Contabilidad General, que forma parte del Poder Legislativo de los EE.UU. y directamente reporta al Congreso. En otros contextos, la credibilidad de la evaluación podría ser derivada de la transparencia del proceso antes que de una ausencia de prejuicios, pero cualquiera que sea la estrategia es fundamental que la evaluación se lleve a cabo de manera que sea percibida por el público en general y las partes interesadas como creíble.

De hecho, incluso en contextos menos politizados que la Oficina General de Contabilidad, la naturaleza política de la evaluación, plantea desafíos para la independencia, la credibilidad y la utilización de las evaluaciones. Y es por eso que la evaluación es mucho más que tratar con metodologías y técnicas que nos permiten adquirir una evidencia lo suficiente buena de lo que ha ocurrido en un proyecto concreto, programa o política. O, como me gusta decir, la metodología de la evaluación tiene que ir mucho más allá de lo que se entiende como la investigación en ciencias sociales, incluyendo la mayor parte de lo que se llama investigación social aplicada. Necesitamos formas específicas para pensar y llevar a cabo las evaluaciones que son exclusivas a la evaluación e incluidas en su naturaleza política y, en su propósito -de superación, de rendición de cuentas y / o iluminación (*enlightenment*).

La naturaleza política de la evaluación significa que, además de proporcionar evidencia sistemática y rigurosa, ésta se eleva, se debe adaptar, y debe responder a y desde un contexto particular para ser creíble y utilizable. Esta naturaleza política también debe reconocerse en las evaluaciones independientes. Aunque parezcan ser apolíticas, la evaluación es siempre e inevitablemente política, ya que se forma a partir de la interacción de los actores con diferentes intereses (algunos de ellos en aparente contradicción, pero todos ellos legítimos),

perspectivas y posiciones de poder. Existe interdependencia entre los actores, y ellos operan en un territorio y período de tiempo concreto. Cualquier proyecto, programa o política que puede ser objeto de evaluación se puede describir en este contexto.

Esta importante presencia de diferentes actores y grupos de interés nos lleva al siguiente punto.

2. ¿Quién participa y tiene voz y voto en la evaluación? Una cuestión clave para la credibilidad

La credibilidad también está inevitablemente ligada a la participación de los interesados en los procesos de evaluación. Un enfoque participativo de la evaluación, además de los beneficios de integración, promueve la apropiación de los interesados de los procesos y resultados de la evaluación. En consecuencia, se plantea también la percepción de la credibilidad de los procesos y resultados de los que han participado y sienten que tienen algo que decir en la evaluación.³ Por supuesto, esto es más fácil decirlo que hacerlo y la participación real de los interesados es un proceso que requiere tiempo y sensibilidad política de los comisionados y los equipos de evaluación. Conocer y comprender el contexto y el programa a evaluar, identificar las principales partes interesadas y sus necesidades de información, incluyendo esas necesidades en las preguntas de evaluación y contrastando o incluso negociando con las partes interesadas, la recopilación de información sistemática y la evidencia de los interesados y de otras fuentes, y contrastando y / o elaborando conclusiones y recomendaciones en colaboración, son algunas de las formas en que los interesados pueden participar. Eso ayudará a que la evaluación sea percibida como más creíble, ya que el proceso ha sido transparente y los interesados han tenido la oportunidad de contribuir a varias partes del mismo.

A menudo, el contexto en el que la evaluación se lleva a cabo, o las limitaciones de tiempo y recursos y de los términos de referencia, no permiten la participación de grupos de interés real en las diferentes partes del proceso. Pero debemos ser conscientes de que para mejorar la credibilidad, es necesario diseñar metodológicamente un proceso de evaluación en el que sea posible la participación de los interesados en algunas partes del proceso de evaluación.

Hasta ahora, hemos discutido la participación de las partes interesadas como requisito metodológico para mejorar la apropiación y por lo tanto, la credibilidad de la evaluación. Pero hay otras cuestiones relativas a la justificación de la participación de los grupos de interés que van más allá de este punto de vista más bien utilitario y pragmático, que también son cruciales para la credibilidad de la misma. Muchos teóricos de la evaluación recuerdan a los evaluadores la centralidad de la participación de los interesados. Por ejemplo, Monnier habla de la

³ La norma 10.1 de las "Normas de evaluación en el sistema de las Naciones Unidas", publicado por UNEG en 2005, dice lo siguiente: "La transparencia y la consulta con las principales partes interesadas son características esenciales en todas las etapas del proceso de evaluación. Esto mejora la credibilidad y la calidad de la evaluación. Puede facilitar la creación de consenso y apropiación de los resultados, conclusiones y recomendaciones". Ver www.uneval.org/documentdownload?doc_id=21&file_id=562

importancia de la "utilidad social" de la evaluación. Que la utilidad social sólo puede alcanzarse con la participación de todos los actores involucrados en el programa o política pública que está siendo evaluada. La evaluación, dice, no sólo debe tener legitimación institucional, técnica o científica, debe tener, ante todo, legitimación política, que viene dada por la participación de los interesados. Sin embargo, pasar por las propuestas de participación que buscan sin verdadera intención el maximizar la participación podría conducir a favorecer a los que retienen las palancas del poder. El modelo de evaluación democrática deliberativa (House & Howe, 2000), por ejemplo, propone procedimientos que aseguran que los desfavorecidos son respetuosamente incluidos en el proceso de evaluación. Para House, la evaluación democrática deliberativa "aspira llegar a conclusiones imparciales teniendo en cuenta todos los intereses, valores y perspectivas, al participar en un amplio diálogo con los principales interesados, y mediante la promoción de una amplia deliberación acerca de las conclusiones del estudio, además de emplear metodologías tradicionales de evaluación "(House, 2005). Este reconocimiento de las desigualdades estructurales y sistémicas que producen, de forma predeterminada, sesgos o preferencias por los favorecidos o poderosos, también es una razón clave de la necesidad de una perspectiva evaluativa basada en la equidad, sensible a temas de género y derechos humanos. Todos estos enfoques nos recuerdan que hay una necesidad de hacer un esfuerzo deliberado para nivelar el campo de juego permitiendo a los más débiles tener una voz en el proceso de evaluación.

La participación de los interesados no es ni simple ni sencilla. Nuevas formas de elaboración de políticas y nuevos instrumentos son necesarios para hacer frente a nuevos y complejos problemas públicos y a los desafíos de política a nivel local, así como a nivel nacional, regional y mundial. Tales complejidades incrementan las posibilidades de participación democrática y el número de interacciones entre los interesados. Los gobiernos se han movido desde una forma centralizada y jerárquica, de arriba hacia abajo, a una forma de gobernanza que implica múltiples tipos descentralizados y controvertidos de actores públicos y privados. Para ser creíble en este entorno cambiante, la práctica y la teoría de la evaluación no puede ignorar las nuevas realidades y ofrecer diseños de evaluación sencillos para situaciones e intervenciones complejas. Otra perspectiva importante con respecto a la credibilidad y la participación de los interesados es la perspectiva de la ciudadanía. En la democracia, como dice Tarja Cronberg (Parlamentario Europeo y anfitrión del evento de la EES "Evaluación en Democracia en el Parlamento Europeo"), la evaluación puede ser una herramienta para el empoderamiento de los ciudadanos. Puede ser también una forma de promover el aprendizaje social, la identificación de las intervenciones políticas prioritarias y la reducción del déficit democrático, en este caso, en la Unión Europea. Al mismo tiempo, los ciudadanos deben ser capaces de confiar en las instituciones para que la evaluación pueda desempeñar un papel. Sin embargo, si se planea una evaluación, se implementa y se utiliza correctamente desde la perspectiva del ciudadano, debería ayudar también a recuperar y construir la confianza institucional.

3. La credibilidad no sólo depende de la calidad e independencia de los evaluadores, sino también de las instituciones y los sistemas donde se conciben, planifican y administran las evaluaciones

Los evaluadores no desempeñan un papel solitario en la evaluación. Son contratados por los comisionados que establecen términos de referencia y clientes que normalmente tienen algo que decir sobre cómo la evaluación debe ser concebida y realizada. Pero los comisionados y los clientes no funcionan solos tampoco. Juegan un papel en las organizaciones, que tienen un propósito, una forma de trabajo y a veces objetivos estratégicos en su trabajo de evaluación. Las evaluaciones pueden realizarse mediante sistemas de evaluación o procedimientos establecidos, por ejemplo, o desarrollando términos de referencia, de búsqueda y contratación de los evaluadores. Por lo tanto, la credibilidad no sólo depende de la calidad e independencia de los evaluadores, sino también de las instituciones y los sistemas donde se conciben, planean y administran las evaluaciones.

Sin embargo, la comunidad de evaluación tiene una tendencia a pensar – a veces exclusivamente- desde el punto de vista del evaluador. Un ejemplo típico es el conjunto de normas, principios y directrices para la conducta ética, que algunas sociedades de evaluación han adoptado en las últimas dos décadas. La mayoría, incluyendo el Comité Conjunto sobre Normas para la Evaluación Educativa, que estableció el primer código de evaluación, sólo parece pensar desde el punto de vista del evaluador. Ellos recomiendan lo que los evaluadores deben hacer para llevar a cabo una evaluación buena, útil, correcta y ética. Incluso las normas para la evaluación en el sistema de las Naciones Unidas, que fueron establecidas por UNEG en 2005 y tienen una perspectiva del sistema de las Naciones Unidas en la mayoría de las secciones (definición, responsabilidad de evaluación, políticas, intencionalidad, imparcialidad, independencia, evaluabilidad, calidad, competencias de la evaluación, transparencia y consulta, ética de la evaluación, seguimiento y contribución a la construcción de conocimiento), confía las dos normas en la sección de calidad de evaluación y las cinco de la ética de la evaluación, en manos de los evaluadores. Las Pautas de Evaluación de la Asociación Británica para Buenas Prácticas en Evaluación⁴ son excepcionales en este sentido, como también lo son para los comisionados, la autoevaluación y para los participantes de la evaluación, así como para los evaluadores.

En suma, hemos pensado más sobre lo que los evaluadores pueden hacer para mejorar la credibilidad de la evaluación, pero menos en cómo las instituciones y los comisionados deben promover la credibilidad. ¿Qué deben hacer las instituciones o los gobiernos? Yo sugeriría que tengan una política clara sobre la evaluación, sobre su propósito, y a quién debería beneficiar. Por ejemplo, un entorno propicio para la

⁴ Ver:

<https://www.evaluation.org.uk/assets/UKES%20Guidelines%20for%20Good%20Practice%20January%202013.pdf>

participación de los actores es más frecuentemente facilitado – u obstruido – por organizaciones, clientes y comisionados que por los evaluadores.

Los instrumentos y mecanismos para la implementación de una política clara de evaluación, que en última instancia mejora la credibilidad, pueden variar enormemente y dependerán de las características del sistema político y del contexto general y la cultura de cada país, entre otras cosas. Para algunos países de Europa, la creación de una institución específica para las políticas públicas de evaluación para todos los sectores podría funcionar (España). En otros lugares, la meta se puede perseguir mejor a través de los avances en cada sector político (Francia). Algunos podrían encontrar que aprobar la legislación concreta que la evaluación requiere es particularmente útil, mientras que otros entienden que este mecanismo podría tener efectos no deseados, como la excesiva burocratización o simplemente no implementarse (Italia). Algunos sistemas de evaluación se basan en el Poder Legislativo (Suiza), mientras que el Poder Ejecutivo es responsable en algunos países (Suecia). Cada sistema político y administrativo, cada situación política, puede requerir una solución diferente para promover la credibilidad.

España ofrece un ejemplo ilustrativo. La *Agencia Estatal de Evaluación de Políticas Públicas y Calidad de los Servicios-AEVAL* se estableció en 2006, después de un proyecto muy bien considerado en el que participaron diversos comisionados expertos. En la primera parte del proyecto en el año 2004, los expertos piden a una agencia, dependiente del Poder Legislativo para tener mejores condiciones para la independencia y por lo tanto, credibilidad. El pedido fue hecho en el entendido que las dificultades que tendría su implementación en el sistema político no presidencial de España, que está formado por las listas electorales cerradas y demanda una fuerte disciplina partidaria en el sistema parlamentario. Finalmente, debido a las dificultades formales que esta solicitud requeriría, la recomendación fue comenzar la Agencia a nivel ejecutivo y eventualmente cambiarlo para depender del Senado cuando una reforma constitucional prevista del Senado estaba por venir. Lo que no se sabía en ese momento es que España estaba a punto de entrar en un período de grave crisis económica, que fue precedido por una crisis política e institucional que comenzó con una tremenda – incluso histórica – confrontación entre los dos partidos principales, que se reflejó especialmente en el Parlamento. Tener a AEVAL dependiendo del Poder Legislativo en este contexto habría tenido consecuencias imprevistas sobre la credibilidad por parte de los administradores y la ciudadanía. Desde mi perspectiva, AEVAL tuvo una mayor oportunidad de mejorar su credibilidad e independencia desde el Ejecutivo que si hubiera estado en el Poder Legislativo.

Desafortunadamente, AEVAL nació en el inicio de una profunda crisis económica. Recortes presupuestarios dramáticos y su reducción gradual como una prioridad política significó que no se desarrolló como se había planeado. En cuanto a su credibilidad, AEVAL comenzó a servir a la administración pública en el comienzo de una grave crisis política, que condujo a una falta de apoyo político cada vez más profundo y sustancial por parte del público español, acompañado por una amplia desconfianza política e institucional (la insatisfacción general por el funcionamiento de la democracia ha continuado en aumento de 42 por ciento en 2007 a 62 por ciento en

2011).⁵ En mi opinión, habiendo llegado a este punto, la única manera de recuperar la confianza institucional y política en España será dando voz verdadera a la ciudadanía. Y la evaluación podría desempeñar un papel importante en esa recuperación.

4. Creación de capacidades en Evaluación tanto para los evaluadores como para los comisionados

Es evidente que credibilidad de evaluación depende en gran medida de la experiencia de los evaluadores – y la de los comisionados. Esta es la razón por la cual cursos de capacitación específica en evaluación, buenos programas de entrenamiento y pasantías, son tan importantes para el fortalecimiento de capacidades de evaluación.

Sin embargo, existen pocos programas de postgrado específicos en evaluación, ya sea en Europa,⁶ en América del Norte o en otras regiones del mundo. Después de muchos años dedicados a la formación del evaluador, estoy convencida que tomar un curso de métodos generales de evaluación, de investigación, de gestión, de políticas públicas o maestrías o doctorados en Desarrollo, con frecuencia no es suficiente para considerar a alguien preparado para ser un buen evaluador o comisionado. Tampoco lo son los cursos de formación profesional de 20 o 40 horas. Una parte muy importante de todo buen entrenamiento en evaluación debería ser 'aprender haciendo', que es la razón crucial para proporcionar *internados* y experiencias prácticas para los nuevos evaluadores. Pero previo a este 'aprender haciendo', también hay un amplio cuerpo de conocimientos y teorías de evaluación, prácticas y disposiciones que forman parte de un plan de estudios de evaluación, y que puede y debe ser enseñado.

En relación con la creación de capacidades, es necesario mencionar los esfuerzos de que la comunidad de evaluación ha hecho hacia su profesionalización. Esos esfuerzos incluyen la elaboración de códigos de práctica. Estos son conjuntos de normas, guías y estándares adoptados por las sociedades de evaluación principalmente durante los años 1990 e inicio de los 2000 (véase Bustelo, 2006). Estos códigos comenzaron a establecer unos límites de lo que un evaluador debería hacer como parte de una buena, útil y éticamente correcta evaluación (con la excepción de las directrices de la sociedad de evaluación de Reino Unido antes mencionados, que añaden lineamientos para comisionados y otros). Estos últimos años han visto desarrollarse esos códigos hasta convertirse en un enfoque en competencias o capacidades, es decir, las capacidades que un evaluador debería tener y desarrollar. La Sociedad Europea de

⁵ La confianza pública en las instituciones ha caído enormemente entre 2007 y 2011. Aunque no tenemos aún los datos actualizados, probablemente ha seguido cayendo en 2013, dado los últimos casos de corrupción y, en mi opinión, el mal manejo de estos casos por el gobierno español. Como ejemplos, aquí están los Promedios en confianza para el 2007 y 2011 para diferentes instituciones: el Rey, de 7.20 a 5,79; el Tribunal Constitucional 6.08 a 5.00; el Congreso de los diputados de 5,54 a 4.87; el Senado de 5,45 a 4.07; el Gobierno Español de 5.46 a 3.93; la UE de 6.67 a 5,79; partidos políticos de 4,22 a 3,38; y los sindicatos 4,58 a 3,26. Datos de CIS (*Centro de Investigaciones Sociológicas*) y de la presentación del prof. Francisco Llera en el 6º Seminario de Verano de Harvard en Investigación Sociológica y Política, Agosto de 2013.

⁶ Ver

www.europeanevaluation.org/images/file/Study_programmes/16_profiles_November%202012.pdf

Evaluación (EES, por sus siglas en inglés) desarrolló un marco de capacidades y lo validó a través de una encuesta entre sus miembros.⁷ Este mismo esfuerzo se ha convertido recientemente en una discusión sobre la posible designación o acreditación de evaluadores.

Canadá ya ha implementado un sistema de acreditación y la Asociación Internacional de Evaluación para el Desarrollo (IDEAS) ha iniciado su propia revisión. Mientras que los temas involucrados son controversiales y sensibles, la Sociedad Europea de Evaluación considera que es el momento adecuado para entablar un debate público sobre los pros y contras de un potencial sistema de revisión entre pares dirigido a la designación y desarrollo profesional. Gracias a Evalpartners, que ha proporcionado fondos para que la Sociedad Europea de Evaluación y la Sociedad de Evaluación de Reino Unido auspicien un taller conjunto diseñado para debatir sobre el potencial de un Sistema Voluntario de Pares Evaluadores (VEPR por sus siglas en inglés), se ha puesto en marcha una discusión en la membresía de la Sociedad Europea de Evaluación. La propuesta de VEPR ha sido diseñada bajo los auspicios del Grupo de Trabajo de Profesionalización Temática (TWG, por sus siglas en inglés) de la Sociedad Europea de Evaluación y la directiva de la Sociedad unánimemente acordó que la propuesta de TWG merece ser expuesta a una amplia audiencia de evaluación como primer paso en un proceso consultivo proactivo.

¿Por qué estoy enlazando estos esfuerzos de profesionalización al fortalecimiento de capacidades de evaluación necesaria como forma de mejorar la credibilidad? Porque estos esfuerzos han sido y seguirán siendo una ayuda invaluable para la construcción de capacidades en evaluación, identificando lo que los evaluadores deben saber, las capacidades en que deberían ser entrenados y eventualmente evaluados para su designación o acreditación. La credibilidad de la evaluación también debe promoverse con un sistema de evaluadores y comisionados 'acreditados'. La calidad y credibilidad de una evaluación no depende sólo de los evaluadores, también estas capacidades deberían ser consideradas para comisionados y unidades de evaluación en otros lugares. La credibilidad también depende de los sistemas, cultura, instituciones y contextos políticos. Cualquier marco o sistema que ayuda a definir cómo debería ser una evaluación creíble, no debe nunca imponerse, debe ser objeto de revisión periódica y de negociación en la comunidad y generalmente debe ser lo suficientemente flexible para ser útil en diferentes contextos.

Desde mi perspectiva, un indicador del grado de consolidación en la función de evaluación es el hecho que los comisionados y administradores reconozcan que ellos también necesitan aprender sobre evaluación y buscar algún tipo de entrenamiento. Una de las mejores características de un curso de evaluación es estar sentados juntos intercambiando y aprendiendo al menos estos dos diferentes roles y perspectivas – evaluadores, y jefes y comisionados de evaluación- unos de otros. Sin embargo, se ha pensado más en la posibilidad de entrenar algunas capacidades y habilidades para

⁷ www.europeanevaluation.org/files/EES_Capabilities_Survey_Evaluation.pdf

diferentes roles en evaluación. Esta es una de las razones por las que creo que deberíamos mirar el cuadro de Desarrollo de Capacidades en Evaluación desde diferentes ángulos y perspectivas. Como un ejemplo ilustrativo, he visto comisionados tan entusiastas en su trabajo de evaluación que han malinterpretado su papel y han dicho al equipo evaluador exactamente cómo debe llevar a cabo la evaluación. Pero el trabajo de un buen comisionado no es sobre la elaboración de Términos de Referencia (como desafortunadamente he visto más de una vez) en los cuales se pasa mucho tiempo explicando la metodología concreta y los métodos que se utilizarán. Todos sabemos que los equipos de evaluación necesitan un marco y contexto bien explícitos, pero también requieren de suficiente libertad para poder hacer un buen trabajo.

5. Credibilidad y metodología: ¿Qué cuenta como evidencia creíble y para quién?

La credibilidad de la Evaluación está relacionada con los datos que deben ser recogidos empíricamente para responder a preguntas de evaluación. Esto tiene un primer nivel, que está relacionado con el ámbito de la información que debe ser recolectada y la fiabilidad de la información adquirida por una evaluación. La cantidad y la fiabilidad de la información, junto con la ausencia de sesgos, es lo que obtenemos a través de terceras personas o de fuentes secundarias. ¿De quién hemos obtenido información? ¿Hemos obtenido información de todo el espectro de partes interesadas, o sólo de algunos de ellos? También está el tema de la honestidad y la imparcialidad, porque la credibilidad requiere que las evaluaciones reporten tanto los éxitos como los fracasos.

La credibilidad está inevitablemente relacionada a la perspectiva metodológica o como la evidencia empírica es obtenida y analizada. Esta relacionada a métodos (por ejemplo, estudios de caso, encuestas, cuestionarios y experimentos). Pero ella siempre está relacionada a una pregunta más filosófica de investigación social, sobre la naturaleza de la fiabilidad sobre que constituye conocimiento y como él es creado, esto es, acerca de epistemología y ontología. Por otra parte, lo que se considera evidencia creíble está claramente mediada por la filosofía fundamental de las nociones de ciencia, como el concepto de paradigma. El debate sobre lo que se considera evidencia creíble proviene de un viejo y recurrente debate sobre la mejor manera de estudiar los fenómenos sociales: el debate cuantitativo-cualitativo.

En el campo de la evaluación, la pregunta acerca de lo que constituye evidencia creíble utilizado para apoyar las alegaciones relativas al impacto de una práctica, programa o política, han vuelto a aparecer con fiereza (*fiercely*) cuando algunas organizaciones internacionales, redes y departamentos federales de EE.UU. han identificado el ensayo controlado aleatorio (RCT) como el "diseño" estándar de oro "para la generación de" evidencia científica "de los programas o políticas evaluadas.

Esto creó mucho desconcierto en la comunidad de evaluación durante la última década, y generó respuestas de varias sociedades de evaluación, incluyendo la Asociación Americana de Evaluación (AEA) en 2003 y la Sociedad Europea de Evaluación de 2007⁸. En 2006, la Universidad de Claremont organizó un simposio sobre

⁸ Consulte la declaración de EES sobre la diversidad metodológica en [www.europeanevaluation.org / library.htm](http://www.europeanevaluation.org/library.htm)

"¿Qué cuenta como evidencia creíble en la investigación aplicada y la práctica de la evaluación?", en la que se invitó a participar a conocidos académicos de la evaluación, tanto en el grupo experimental-cuantitativo y los métodos no-positivista-constructivistas y cualitativos. Ese simposio condujo a la producción de un volumen editado con el mismo título que el simposio (Donaldson, Christie & Mark, 2009), que se organiza en torno a los paradigmas de investigación social como un marco para el debate sobre la evidencia creíble. Aunque se podría argumentar que la misma diferenciación entre los enfoques experimentales y no experimentales es de alguna manera normativa y tiende hacia la condición definida -la definición de "el otro" por la ausencia de esa condición - esta fue la primera vez que la credibilidad y los conceptos de evidencias e impacto se debatieron abiertamente y en profundidad desde una perspectiva metodológicamente plural, y no exclusivamente relacionada con una postura epistemológica concreta, como el experimentalismo.

Desde mi punto de vista, no puede ser de otra manera, ya que la disciplina de la evaluación ha evolucionado desde perspectivas comunes y múltiples métodos y enfoques, lo que permite el debate entre los evaluadores que vienen de tradiciones de estudio muy diferentes.

Debido a su carácter práctico y aplicado, y la necesidad de evidencia creíble desde diferentes perspectivas para responder a diversas preguntas en diferentes contextos, la evaluación ha sido uno de los primeros ámbitos en los cuales investigadores y evaluadores cualitativos y cuantitativos han intercambiado puntos de vista, establecido redes y hablado uno con otro. Por otra parte, la comunidad de evaluación estuvo rápidamente dispuesta a abrazar el enfoque de los métodos mixtos.

En esta "era de pluralismo paradigmático" (Greene, 2013: 111) y el entendimiento necesario entre las diferentes perspectivas sobre la investigación social, este enfoque de métodos mixtos es ampliamente aceptado en la comunidad de evaluación. Como Donna Mertens y Sharlene Hesse-Biber dice en la Nota del Editor, en su reciente volumen, *Métodos mixtos y la credibilidad de las pruebas de evaluación*, en nuevas direcciones para la evaluación:

*Un viejo proverbio italiano dice: "Lo que es viejo es nuevo, lo que es nuevo es viejo" (Melfi, 2011). Esta cita caracteriza la historia de los métodos mixtos en la comunidad de evaluación en cuanto a que los métodos mixtos han sido utilizados por los evaluadores durante muchos años. Muchos evaluadores intuitivamente llegaron a la conclusión de que las evaluaciones de los programas sociales complejas se podrían mejorar mediante el uso de varios métodos, por lo que la combinación de datos cuantitativos y cualitativos en el mismo estudio no es nada nuevo. La atención a los métodos mixtos de evaluación fue evidente en las nuevas orientaciones para la evaluación (NDE) editado por Jennifer Greene y Valerie Caracelli en 1997 (Greene y Caracelli, 1997). Desde entonces, la atención a los métodos mixtos se ha incrementado exponencialmente, como lo demuestra el lanzamiento de la Revista de Métodos de Investigación Mixto en 2007, que tenía un factor de impacto inicial de 2.219 y ocupaba el quinto lugar de los 83 revistas de las ciencias sociales, la categoría interdisciplinaria, de acuerdo al Journal Citation Reports 2010 por Thomson Reuters (2011). El Topical Interest Group (TIG) de la Asociación de Evaluación Americana (AEA): **Métodos mixtos de evaluación** fue fundada en*

2010 y rápidamente se convirtió en uno de los más grandes TIGs de la AEA. Y, el Manual Sage de Métodos mixtos en la Investigación Social e Investigación del Comportamiento (Tashakkori y Teddlie, 2010) se encuentra en su segunda edición. (Mertens y Hesse-Biber, 2013: 1)

Esta importante pieza de trabajo tiene como objetivo preguntar "sobre la conexión entre la convocatoria por programas basados en evidencias y la contribución potencial de los métodos mixtos para la creación de pruebas creíbles. El propósito de este tema es examinar las contribuciones del método de evaluación mixta y sus filosofías emergentes, las teorías y las prácticas (...), así como abrir la posibilidad de aumentar la credibilidad de las evaluaciones que parten de varias posturas paradigmáticas, como postpositivismo, el pragmatismo, el constructivismo, y transformativismo" (Mertens y Hesse-Biber, 2013: 3).

El enfoque de los métodos reclama que el uso mixto no sólo debe funcionar al nivel de métodos, sino también en la metodología y el nivel de la epistemología. Como Mertens y Hesse-Biber dicen, "es importante entender que los métodos mixtos no se trata sólo de mezclar y combinar () métodos. El uso de cualquier método o conjunto de métodos en una evaluación también está estrechamente vinculado a epistemologías, metodologías específicas (perspectivas teóricas), y supuestos axiológicos, así como estar conectado a particulares perspectivas de los interesados "(Mertens y Hesse-Biber, 2013: 5-6).

Este volumen plantea la necesidad de ser consciente y explícito acerca de las perspectivas teóricas y los supuestos. Jennifer Greene afirma que evaluadores con métodos mixtos deben ser "explícitos sobre los supuestos paradigmáticos que enmarcan y orientar su trabajo", y que "una cuidadosa explicación de lo que se mezcla en un estudio de métodos mixtos contribuye a la subsecuente garantía de la credibilidad de los resultados". Para ella, "es una responsabilidad fundamental del investigador" hacer suposiciones explícitas, tales como la "naturaleza del mundo social, lo que se considera como conocimiento garantizado, metodología defendible, y el papel de la investigación social en la sociedad". Ellos también deben "justificar los valores que ellos invocan -los valores de la distancia, la participación, la inclusión, objetividad, generalidad, contextualidad, la acción social", y así sucesivamente. Esto es particularmente importante en contextos de evaluación, debido a que están saturados con valores "(Greene, 2013: 111-112).

En resumen, la credibilidad, la evidencia y el impacto no son conceptos exclusivamente válidos para posiciones positivistas, por lo que deben ser exploradas y definidas por otras perspectivas paradigmáticas. El positivismo ha sido el paradigma dominante durante muchos años, pero no es necesariamente el caso único -como el pluralismo metodológico y paradigmático en la comunidad de evaluación ha demostrado. Los evaluadores de métodos mixtos proponen avanzar en el debate de pruebas creíbles, haciendo explícitos los valores, así como las opciones ontológicas, epistemológicas y metodológicas. Así que se necesita transparencia paradigmática y metodológica para la credibilidad. En mi opinión, debemos insistir en que esta transparencia se ejerce no sólo por los evaluadores de métodos mixtos, sino por todos los evaluadores. En aras de la credibilidad, creo que entendimientos "clásicos" y dominantes, como el experimentalismo, no deben darse por sentados, y las opciones paradigmáticas y

metodológicas que impulsan esas perspectivas deben ser explicadas y hacerse explícitas. Esto sería un verdadero reconocimiento de que hay otros modos de investigación que no son jerárquicamente inferiores. De esta manera, el pluralismo metodológico se hace real - ya no será la "norma" una sola perspectiva, mientras que otras opciones "alternativas" tienen que ser justificadas.

6. Comunicación y presentación de informes: otro tema clave para la credibilidad de la evaluación.

Como punto final importante en este debate sobre la credibilidad de la evaluación, me gustaría hablar de la comunicación y presentación de informes. A pesar de la falta de tiempo y espacio en este documento guía, que no me permite desarrollar aún más este punto, me gustaría mencionar al menos tres aspectos o preguntas para la discusión en la comunidad de práctica en línea. Estas son: la cuestión de la transparencia y lo que debe hacerse explícito en el informe, la necesidad de una comunicación fluida con las partes interesadas durante el proceso, y la accesibilidad de los informes de evaluación.

Si la evaluación es para ser creíble, debe ser transparente y explicar los contextos políticos e institucionales en los que se llevó a cabo. También se debe tener en cuenta - e incluir en el informe de evaluación - temas paradigmáticos, metodológicos y los valores relacionados.

La comunicación fluida con los comisionados y las partes interesadas durante el proceso de evaluación es conocido que promueven un mejor uso de los resultados de la evaluación (véase, por ejemplo, Torres, Preskill y Piontek, 1996). Pero la comunicación fluida también es importante para mejorar la credibilidad del proceso de evaluación y la evidencia que resulta de ella. Si los comisionados y los interesados tienen información periódica y transparente durante el proceso de evaluación, esto se traduce en una mayor probabilidad de credibilidad percibida por ellos.

Una cuestión clave para la transparencia y la credibilidad está relacionada con la accesibilidad a los informes de evaluación. El proceso de evaluación pudo haber sido bien diseñado, la evaluación pudo haber sido realizada con honestidad y transparencia, pero si el informe final no es accesible al público será casi imposible que sea percibida como creíble por los que no tienen acceso a ella. La facilidad de acceso a la información a través de Internet ha mejorado enormemente la apertura, pero también desafía los casos en los que no hay acceso inmediato a los informes de evaluación.

Referencias Bibliográficas

Bustelo, M. 2006. "The Potential Role of Standards and Guidelines in the development of an Evaluation Culture in Spain" *Evaluation*, Vol. 12(4) pp.436-452.

Chelimsky, E. 1998. "The role of experience in formulating theories of evaluation practices". *American Journal of Evaluation*, 19(1): 35-55

- Donaldson, S., Christie, C. & Mark, M. 2009. *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* Thousand Oaks, CA: Sage Publications.
- Greene, J. C. (2013). Reflections and ruminations. In D. M. Mertens & S. Hesse-Biber (Eds.), *Mixed methods and credibility of evidence in evaluation. New Directions for Evaluation*, 138, 109–119.
- House, E. 2005. “Deliberative Democratic Evaluation”. In S. Mathison (Ed), *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Publications.
- House, E. & Howe, K. 2000. “Deliberative Democratic Evaluation in Practice”. In George F. Madaus, Michael S. Scriven, Daniel L. Stufflebeam (eds.), *Evaluation models : viewpoints on educational and human services evaluation*. Boston : Kluwer-Nijhoff.
- Mertens, D. M., & Hesse-Biber, S. (2013). Mixed methods and credibility of evidence in evaluation. In D. M. Mertens & S. Hesse-Biber (Eds.), *Mixed methods and credibility of evidence in evaluation. New Directions for Evaluation*, 138, 5–13.
- Monnier, E. 1995. *Evaluación de la acción de los poderes públicos*. Madrid, Instituto de Estudios Fiscales. Ministerio de Economía y Hacienda.
- Torres, R., H. Preskill, and M. Piontek. 1996. *Evaluation Strategies for Communicating and Reporting*. Thousand Oaks: Sage