

# RQ+



**Calidad de Investigación Plus**

---

## Un Enfoque Integral para Evaluar la Investigación

### **Zenda Ofir**

Evaluadora independiente  
Profesora honoraria, Universidad de Stellenbosch, Sudáfrica

### **Thomas Schwandt**

Profesor, Universidad de Illinois, Urbana-Champaign, Estados Unidos

### **Colleen Duggan**

Especialista de programas sénior, División de Políticas y Evaluación,  
Centro Internacional de Investigaciones para el Desarrollo (IDRC), Canadá

### **Robert McLean**

Especialista de programas sénior, División de Políticas y Evaluación,  
Centro Internacional de Investigaciones para el Desarrollo (IDRC), Canadá  
Evaluador líder, Institutos Canadienses de Investigación sobre Salud, Canadá

## Agradecimientos

Los autores agradecen la revisión crítica de su trabajo llevada a cabo por colegas de todo el mundo. Los comentarios, las correcciones y los consejos de los siguientes pares nos han permitido pulir nuestras ideas y han ampliado nuestro punto de vista. Cualquier error en los datos o los razonamientos es responsabilidad exclusiva de los autores.

**Sarah Bowen**

Universidad de Alberta, Canadá

**Juliana Bravo**

World University Service, Canadá

**Fred Carden**

Using Evidence Inc, Canadá

**Fanie Cloete**

Universidad de Johannesburgo, Sudáfrica

**Leandro Echt**

Politics & Ideas, Argentina

**Amy Etherington**

IDRC, Canadá

**Oswaldo Feinstein**

Universidad Complutense de Madrid, España/Argentina

**Trisha Greenhalgh**

Universidad de Oxford, Reino Unido

**Luize Guimaraes**

Oxford Policy Management, Mozambique

**Brent Herbert-Copley**

Consejo de Investigaciones en Ciencias Sociales y Humanidades de Canadá, Canadá

**Lise Korsten**

Universidad de Pretoria, Sudáfrica

**John Lavis**

Universidad McMaster, Canadá

**Enrique Mendizabal**

OnThinkTanks.org, Perú

**Heidi Monk**

Artisanal Gold Council, Canadá

**Savi Mull**

Red de desarrollo mundial, India

**Claire Nicklin**

Fundación McKnight, Ecuador

**Umezurike Linus Opara**

Universidad Stellenbosch, Sudáfrica

**Andrew Shaw**

Departamento de Desarrollo Internacional, Reino Unido

**Suneeta Singh**

Amaltas, India

**Matthew Smith**

IDRC, Canadá

**Ian Thornton**

UK Collaborative on Development Sciences, Reino Unido

**Cecilia Tortajada**

Universidad Nacional de Singapur, Singapur

**Vanesa Weyrauch**

Politics & Ideas, Argentina

**Tricia Wind**

IDRC, Canadá



# Introducción

**Evaluar la calidad de investigación**<sup>1</sup> no es una idea nueva ni revolucionaria. Desde hace mucho tiempo, los investigadores debaten sobre los mejores criterios y medios para establecer el rigor científico y la importancia de los estudios empíricos en las ciencias naturales, sociales y conductuales. La diferencia es que las inquietudes sobre la calidad de investigación han adquirido un nuevo significado para las entidades académicas, gobiernos, fundaciones, asociaciones sin fines de lucro y organizaciones no gubernamentales e intergubernamentales frente a nuevos intereses globales respecto de niveles de calidad, medición del desempeño, rendición de cuentas, políticas y prácticas basadas en evidencias, y el aprovechamiento de los recursos a la hora de invertir en investigación.

A partir de estas inquietudes globales, se analizaron con detenimiento los puntos fuertes y las limitaciones de los métodos tradicionales que se utilizan actualmente para evaluar el mérito científico de las investigaciones: el deliberativo, como la revisión por pares, y el analítico, como el análisis bibliométrico (Feller y Stern, 2006; Wilsdon *et al.*, 2015). A su vez, ha resurgido el interés sobre la importancia de estudiar la utilización y los procesos de intercambio del conocimiento, y se ha concentrado especial atención en definir y evaluar los impactos o beneficios de la investigación para la sociedad. Esto ha permitido comprender con mayor profundidad que la ciencia y su afán por producir descripciones y explicaciones de los ámbitos sociales y naturales con respaldo empírico ya no pueden considerarse una iniciativa mayormente académica, separada de la preocupación por alcanzar metas sociales.

En el campo de la investigación que recibe financiación internacional y que se centra en las prioridades de desarrollo en países de ingresos bajos a medios, se ha generado un intenso debate e interés sobre la calidad y el uso de la investigación, así como sobre la relación entre el trabajo de investigación y las metas y los valores sociales. Es para esta área —Investigación para el Desarrollo (I4D)<sup>2</sup>, que se diseñó el enfoque para evaluar la calidad de investigación del que trata este informe.

El Centro Internacional de Investigaciones para el Desarrollo (IDRC) en Ottawa, Canadá, financia y facilita principalmente investigaciones para el desarrollo basadas en el Sur Global<sup>3</sup>. Si bien determinar e impulsar la investigación de alta calidad es el núcleo de su trabajo, el Centro

---

<sup>1</sup> Para IDRC, el término "calidad" en "calidad de investigación" se refiere a la calidad del proceso de investigación así como a la calidad de los resultados de la investigación

<sup>2</sup> Este no es el único sitio donde se está dando este debate. Otro ámbito donde resulta muy evidente el creciente interés en estos temas es el sanitario, por ejemplo, Panel on Return on Investment in Health Research (2009) y Guthrie *et al.* (2013).

<sup>3</sup> El mandato del IDRC descrito en la Ley IDRC de 1970 establece que el Centro debe "[...] iniciar, promover, apoyar y conducir investigaciones en torno a los problemas de las regiones en desarrollo del mundo y sobre los medios para aplicar y adaptar el conocimiento científico, técnico y de otro orden con el objetivo de impulsar el progreso social y económico de dichas regiones."

también tiene una importante función evaluadora. Se esfuerza por analizar sus inversiones más significativas utilizando diversos mecanismos, estructuras y plataformas de evaluación, tanto centralizados como descentralizados<sup>4</sup>. Todos los marcos de investigación que emplea deben adaptarse a la heterogeneidad de sus programas, tanto en relación con la diversidad de *temas* que aborda IDRC —como tecnología para la seguridad alimentaria, salud mundial, cambio climático y crecimiento económico inclusivo— como en cuanto a los tipos de *actividades* que apoya —por ejemplo, investigación, fortalecimiento de capacidades, promoción del uso de la investigación para políticas y prácticas, o creación de grupos de apoyo y de redes— y su participación en calidad de par, mentor o intermediario.

En 2012, el IDRC propuso un nuevo enfoque para evaluar la calidad de investigaciones que financia. El resultado se presenta aquí con el nombre de “Marco de Referencia RQ+”. Este trabajo fue motivado por el deseo del IDRC de avanzar en las prácticas globales de evaluación de la investigación y, en lo pragmático, por la necesidad de establecer un grado de normalización y de transparencia para la valoración de la calidad investigativa —esto último constituye una prioridad para sus ejercicios evaluativos de rendición de cuentas, las llamadas ‘Revisiones Externas’<sup>5</sup>.

En este informe se describen en primer lugar los motivos por los que implementamos esta nueva estrategia para evaluar la calidad de investigación. A continuación, se analiza en detalle el instrumento de evaluación en sí mismo, seguido por una sección con lecciones aprendidas a partir del primer intento de implementación del mismo en un conjunto de Revisiones Externas. En resumen, el informe reflexiona sobre los usos potenciales del instrumento y sobre ideas para su futuro mejoramiento.

Lo que se presenta aquí es la versión inicial del Marco de Referencia RQ+, que se encuentra en proceso de revisión por IDRC a fin de ser adaptado a sus circunstancias y usos específicos. Se invita a los lectores —tanto a investigadores que evalúan la calidad de investigación como a quienes financian las investigaciones— a tratar a este Marco de Referencia como una herramienta dinámica y en evolución, que pueden adaptar y modificar para sus propios propósitos.

Esperamos que este enfoque para evaluar la calidad de investigación sea mejorado por los profesionales que lo utilicen y por académicos abocados a la calidad y al uso de la investigación que lo estudien en profundidad.

---

<sup>4</sup> Los detalles completos del enfoque del IDRC para la evaluación se encuentran disponibles en: [idrc.ca/EN/Programs/Evaluation/Pages/default.aspx](http://idrc.ca/EN/Programs/Evaluation/Pages/default.aspx)

<sup>5</sup> Las Revisiones Externas del IDRC tienen cuatro focos de atención: i) coherencia, eficacia y adecuación de las estrategias de implementación de programas; ii) calidad de investigación respaldada por cada programa que se analiza; iii) relevancia de los resultados de los programas, y iv) cuestiones que deben ser sometidas a el Consejo de Gobierno y el equipo de dirección del IDRC.

## PARTE 1

# Fundamentos

Los elementos tradicionales que surgen al evaluar la calidad de investigación están atravesados por dos amplias consideraciones. En primer lugar, durante mucho tiempo se ha creído que diseñar y evaluar la investigación son tareas que conciernen solo a los científicos. Los valores científicos, como la transparencia, la objetividad, la atención a la evidencia empírica y la honestidad intelectual han guiado globalmente el proceso. Además, se establecieron criterios específicos para determinar el rigor del diseño y la implementación de la investigación; la compilación de datos; la confiabilidad y la validez interna y externa de las aseveraciones científicas, y el valor de las teorías científicas en cuanto a su coherencia lógica, replicabilidad, generatividad y su refutabilidad lógica. Para evaluar la calidad de investigación científica a la luz de estos valores y criterios, durante mucho tiempo se ha recurrido a las revisiones por pares, un mecanismo que durante la última década fue cada vez más complementado por el análisis bibliométrico y otros métodos cuantitativos y, en menor medida, por estudios sobre reputación. En líneas generales, este enfoque integral para evaluar la investigación ha regido la manera en que piensan los científicos, más allá del tipo específico de investigación: aplicada, básica, orientada al uso, clínica, de desarrollo o experimental.

En segundo lugar, por lo general, se sostiene que para determinar el mérito científico de la investigación no es necesario evaluar su uso, su aceptación ni su impacto. En otras palabras, la calidad de investigación se suele juzgar exclusivamente según valores y criterios científicos. Podemos decir que este es un pensamiento centrado en la ciencia: evaluar la investigación se reduce a una tarea puramente científica, aislada de su contexto de utilización (McNie *et al.*, 2015). Aunque al evaluar la investigación se analiza cada vez más su impacto más allá de lo académico y se tiene en cuenta el rendimiento de la

inversión (Banco Mundial y Elsevier, 2014; DFID, 2014), estas inquietudes no son, en términos estrictos, ni académicas ni científicas. Según este argumento, esto se debe a que se encuentran en el complejo y subjetivo mundo de la formulación de políticas y la práctica, que se ubica fuera del sistema cerrado de la ciencia.

Esta imagen tradicional de lo que implica evaluar la calidad de investigación está siendo objeto de análisis y de cambios significativos. La comunidad científica se pregunta si la cantidad de citas y el factor de impacto de las publicaciones especializadas son medidas confiables y válidas de la calidad de una investigación (Eyre-Walker y Stoletzki, 2013; Universidad de Gothenburg, 2013). También preocupa que, en entornos académicos, los incentivos al desempeño basados en estas mediciones puedan obstaculizar el tipo de erudición comprometida y que trasciende disciplinas que algunos consideran como un factor crítico para fomentar la innovación. Frente a los temores de que las mediciones bibliométricas para evaluar la investigación sean mal utilizadas o mal interpretadas —y que incluso dañen al sistema de investigación para el que están supuestas a evaluar y mejorar—, se publicó el *Manifiesto de Leiden sobre indicadores de investigación*, un cuerpo de diez principios que orientan el uso responsable de estas mediciones cuantitativas, en la prestigiosa revista científica *Nature* (Hicks *et al.*, 2015).<sup>6</sup> Esto no significa que la comunidad científica desee renunciar a la autoridad y a la dedicación por evaluar el mérito científico de sus investigaciones. No obstante, existen dudas de que los indicadores que se utilizan tradicionalmente para juzgar la calidad de investigación dentro de las organizaciones pertinentes sean los más útiles para evaluar su posible beneficio social (Sarewitz, 2011).

<sup>6</sup> Estos son los diez principios del Manifiesto: (1) la evaluación cuantitativa debe respaldar la evaluación cualitativa realizada por expertos; (2) el desempeño debe ser medido de acuerdo con las misiones de investigación de la organización, el grupo o el investigador; (3) debe protegerse la excelencia en la investigación que tenga relevancia local; (4) los procesos de recopilación y análisis de datos deben ser abiertos, transparentes y simples; (5) debe permitirse a los evaluados que verifiquen los datos y los análisis; (6) deben tenerse en cuenta las diferencias en las prácticas de publicación y citación entre campos científicos; (7) la evaluación de investigadores individuales debe basarse en la valoración cualitativa de su currículum; (8) debe evitarse la minuciosidad impropia y la falsa precisión; (9) deben reconocerse los efectos sistémicos de la evaluación y los indicadores; (10) los indicadores deben ser examinados y actualizados periódicamente.

Asimismo, una creciente cantidad de investigaciones encuentra que el uso de la investigación no se reduce a la difusión y a la traducción, sino que más bien se trata de un proceso complejo, iterativo, interpretativo, interactivo y social que implica relaciones e intercambios entre investigadores, formuladores de políticas y profesionales (Nutley *et al.*, 2007; Lavis *et al.*, 2006; Greenhalgh y Wieringa, 2011; Bowen y Graham, 2015). Se está demostrando, en estudios sobre la sociología de la producción del conocimiento científico, que la visión tradicional de que la ciencia es, en su mayoría, una empresa académica separada de la sociedad está dando paso a diversas interpretaciones de la interrelación entre ambas. En la actualidad, la ciencia es entendida como un proceso social influenciado por valores y normas tanto individuales como sociales. Además de hacerse hincapié en el contexto de aplicación de la ciencia, se ha tomado consciencia de que múltiples actores participan en el proceso de producción del conocimiento y que esto tiene consecuencias para las clases de conocimiento que se producen y en la manera en que se comunican (Nowotny *et al.*, 2003).

Por ejemplo, en la actualidad, la relevancia y la legitimidad que los procesos y productos de investigación presentan para el público no-científico se consideran importantes para comprender la producción de conocimiento científico. Incluso, es probable que se desestime información científica "correcta" que es considerada irrelevante para las necesidades de decisores particulares o que ignora sus puntos de vista y sus valores (Clark *et al.*, 2006). Varios académicos que estudiaron cómo evaluar la investigación en los Países Bajos describen así esta realidad:

*"La mayor parte de las investigaciones se producen en un complejo contexto socioeconómico, donde existen exigencias de varios actores sociales; (...) abordan cuestiones complejas (como el SIDA, el calentamiento global, las migraciones, la identidad cultural); suelen ser multi-, inter- o transdisciplinarias y se llevan a cabo en contextos donde operan expertos con distintas trayectorias, conocimientos y experiencia, y donde deben negociarse diferentes exigencias e intereses. Esta complejidad requiere un enfoque de evaluación diferente de la tradicional revisión por pares, que se enfoca en la excelencia científica y recurre a la publicación en revistas científicas de alto impacto como indicadores principales. En nuestro*

*enfoque, la calidad se define como un concepto multidimensional que incluye la experiencia de las partes interesadas en diferentes dominios sociales"* (Spaapen *et al.*, 2007).

Por último, en la investigación para el desarrollo internacional, se reconoce ampliamente que los proyectos de investigación abordan "problemas retorcidos"; que la labor de investigación está subordinada a circunstancias socioeconómicas, políticas y culturales; que múltiples agentes (incluidas asociaciones y redes) con intereses diferentes participan en la producción de conocimiento, y que los mecanismos que producen impactos a partir de las intervenciones y la propia investigación son variables e impredecibles, y a menudo resistidos. Es decir, que por lo general se entiende que la investigación se evalúa en entornos que dependen en gran medida del contexto y con cierto grado de conflicto. Como se muestra en el recuadro 1, tanto las circunstancias en las cuales se desarrollan las investigaciones financiadas por IDRC como los tipos de investigaciones que apoya constituyen un buen ejemplo del tipo de dependencia del contexto que debe considerarse una influencia clave a la hora de evaluar la calidad de investigación.

La principal consecuencia derivada del discurso actual sobre la valoración de la calidad de investigación es que los investigadores están explorando cómo se puede partir de un interés permanente en los valores científicos e incorporar eficazmente tanto criterios orientados a los usuarios para juzgar la calidad, como la participación de los usuarios de la investigación en el proceso. Por lo tanto, el RQ+ que se describe en la Parte 2 del documento es sustentado por los dos puntos clave a continuación:

## 1

### **Refleja que el mérito científico es una condición necesaria pero insuficiente para juzgar la calidad de investigación.**

La experiencia científica, aunque muy valorada, no es la única experiencia que importa. La evaluación también debe tomar en cuenta si actores involucrados y potenciales usuarios consideran que la investigación es significativa y legítima, y debe prestar singular atención a la eficacia de la investigación en cuanto a su pertinencia, su uso y su impacto. Esta segunda vertiente de la evaluación depende de la experiencia de actores adecuados a la hora de juzgar la legitimidad e importancia del trabajo de investigación.

## 2

**Reconoce que determinar la efectividad de la investigación no es tarea exclusiva de los investigadores, los financiadores ni los directores de los proyectos.** Aspectos como el uso, la influencia y el impacto de la investigación no necesariamente se encuentran dentro de su esfera de control inmediata. No obstante, estos grupos sí desempeñan una función clave al determinar, hasta cierto punto, las condiciones que pueden incrementar tanto el interés de los usuarios en la investigación como el uso que estos hacen de ella. Por lo tanto, es razonable que los investigadores, los directores y los financiadores se hagan responsables del buen posicionamiento de la investigación para su uso.

Estas consideraciones se reflejan en la Figura 1, que destaca que la calidad y la efectividad de la investigación no son conceptos estrechamente vinculados. En la mente de muchos interesados, estos conceptos se superponen y fusionan.

Las tres esferas de la Figura 1 ilustran el proceso de desarrollo de las investigaciones y su progresión no lineal desde la producción hasta el uso. Las esferas muestran que la conceptualización y la ejecución técnicas de la investigación, junto con la producción de resultados, se encuentran mayormente bajo el control de investigadores, financiadores y directores de programas. La influencia que la investigación puede tener sobre las políticas y las prácticas a veces se encuentra en su

### RECUADRO 1 El “contexto” del IDRC para la investigación

#### Atributos de las investigaciones financiadas por el IDRC

- Se trata de investigaciones “inspiradas por el uso”<sup>7</sup> que son relevantes para las políticas/prácticas, abordan problemas, se orientan hacia soluciones o se basan en prioridades nacionales o locales.
- Tienen carácter multi-, inter- o transdisciplinario, que a veces atraviesan campos dispares.
- Utilizan principalmente métodos mixtos.
- Abordan problemas complejos e integradores.
- Son conscientes y respetuosas de las opiniones, el conocimiento y los contextos locales, en especial, en el Sur Global.
- Son conscientes de los riesgos que enfrentan las personas y sociedades vulnerables, así como las organizaciones, sistemas y contextos frágiles.

#### Atributos del apoyo complementario del IDRC a los investigadores

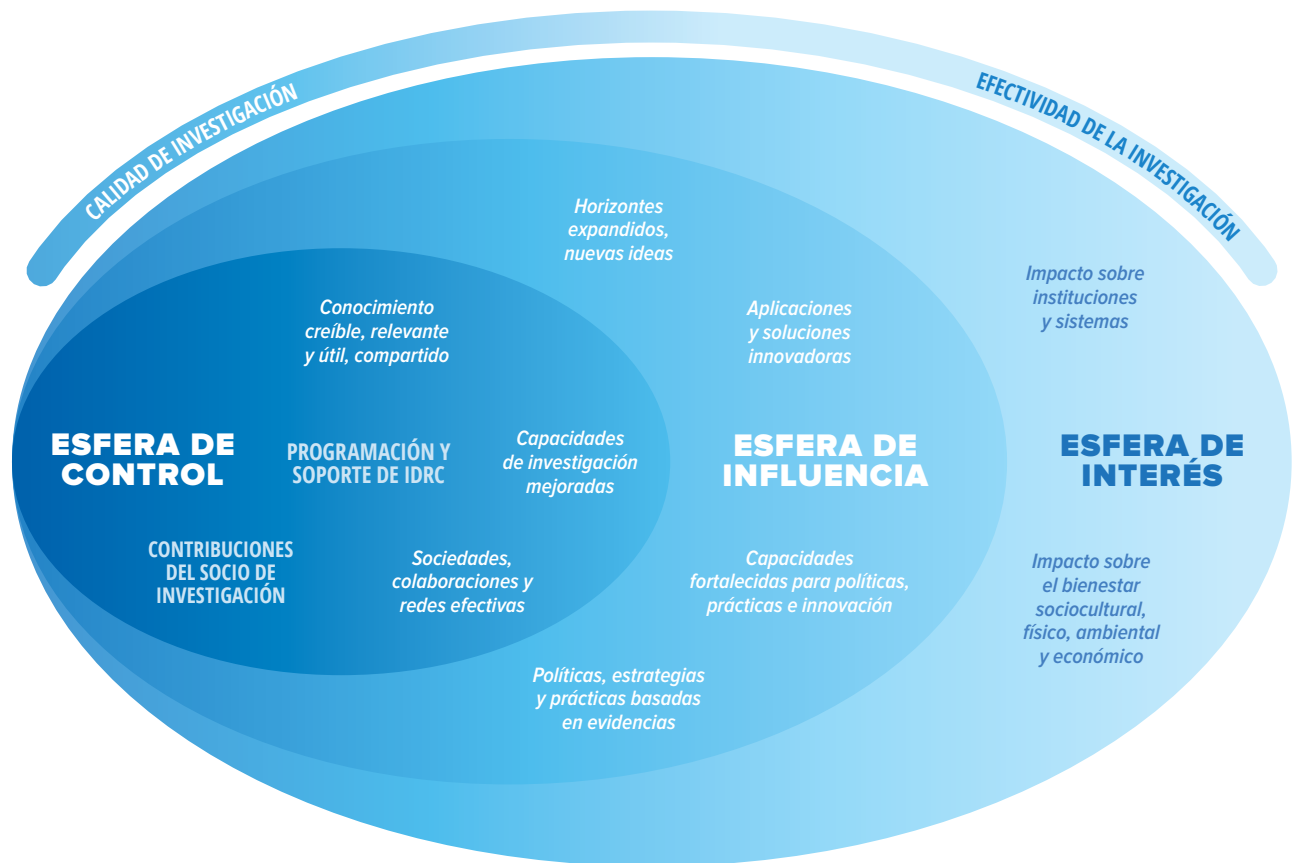
- Fortalece las capacidades de investigación de personas y organizaciones por medio de inversiones de largo plazo.
- Asume riesgos, por ejemplo, al apoyar campos de trabajo completamente nuevos, participar en desafíos regionales o globales complejos y respaldar acciones en entornos conflictivos o asolados por la pobreza.
- Impulsa la generación de conocimientos en el Sur Global.
- Facilita redes de investigación, vínculos entre las investigaciones y las políticas, y el acceso a recursos.
- Construye comunidades y redes que promueven el cambio.
- Se enfoca en efectuar cambios en políticas, prácticas, sistemas institucionales y tecnologías.
- Colabora como mentor, asesor, par e intermediario.

<sup>7</sup> Fuente: Ofir y Schwandt, 2012.



esfera de influencia, pero no siempre. Y, aunque los impactos de largo plazo sobre el desarrollo a los cuales contribuye la investigación pueden encontrarse en su esfera de interés, estos actores, por lo general, tienen muy poca influencia sobre cómo se usan y se transforman los productos de la investigación, y sobre cuándo y cómo se plasmarán estos impactos. Por lo tanto, a la hora de definir, monitorear y evaluar la calidad y la efectividad de la investigación, es importante tener en cuenta que su aceptación, su uso, su influencia y su impacto dependen de la interacción de múltiples actores, organismos y circunstancias sociopolíticas.

**FIGURA 1** Calidad de investigación, efectividad de la investigación y esferas de control



📌 Fuente: Adaptado de Ofir y Schwandt, 2012

# El Marco de Referencia RQ+

# RQ+

**Calidad de Investigación Plus**  
Un Enfoque Integral para  
Evaluar la Investigación

El RQ+ es un método sistemático para definir y evaluar la calidad de investigación y su posicionamiento para el uso y producir impacto. Es adaptable a contextos, valores, mandatos y objetivos, y apoya procesos de planificación, gestión y aprendizaje en cualquier etapa de un proyecto de investigación, programa o portafolio de subvenciones.

## Componentes del Marco de Referencia

El RQ+ contiene tres componentes principales:



### 1. INFLUENCIAS CLAVE

Se identifican las influencias contextuales que probablemente afecten el desempeño investigador, tanto limitadoras como habilitadoras, dentro o fuera del proyecto de investigación.

Al calificar las influencias clave utilizando una escala de tres puntos (por ejemplo, bajo, medio y alto), se establece un perfil de riesgo en el cual se basa la evaluación de calidad.

Las influencias clave pueden ser 1) limitadoras (negativas) o 2) habilitadoras/favorables (positivas).

#### Ejemplos de la experiencia del IDRC:

- 1) Madurez del campo de investigación
- 2) Fortalecimiento de la capacidad para la investigación
- 3) Riesgo en el entorno de datos
- 4) Riesgo en el entorno investigativo
- 5) Riesgo en el entorno político



### 2. DIMENSIONES Y SUBDIMENSIONES

Las cuatro dimensiones y sus subdimensiones engloban los criterios de evaluación de la calidad.

Adaptadas para IDRC:

1. **Integridad de la investigación**
2. **Legitimidad de la investigación**
  - 2.1 Abordaje de consecuencias potencialmente negativas.
  - 2.2 Respuesta ante cuestiones de género
  - 2.3 Inclusividad
  - 2.4 Implicación con el conocimiento local
3. **Importancia de la investigación**
  - 3.1 Originalidad
  - 3.2 Pertinencia
4. **Posicionamiento para el uso**
  - 4.1 Accesibilidad a / intercambio de conocimiento
  - 4.2 Oportunidad (temporal) y usabilidad



### 3. ESCALA DE EVALUACIÓN

El desempeño se describe utilizando rúbricas adaptables para la calidad de investigación.

Se describe cada influencia clave, dimensión y subdimensión utilizando rúbricas específicas que combinan medidas cuantitativas y cualitativas.

Las calificaciones en una escala de 8 puntos muestran cuatro niveles de desempeño (o progreso). Este es un ejemplo. Las escalas deben diseñarse para cumplir un propósito o una intención.



## Hoja de ruta del Marco de Referencia

Los componentes del RQ+ forman una relación dinámica



## PARTE 2

# El Marco de Referencia para la Evaluación RQ+

### Componentes del Marco de Referencia

El Marco de Referencia para la Evaluación RQ+ contiene tres componentes principales: i) influencias clave; ii) dimensiones y subdimensiones de la calidad de investigación, y iii) rúbricas evaluativas. En esta sección se abordan algunos aspectos de estos componentes. Se puede acceder a la totalidad de los componentes —incluidas descripciones detalladas de las influencias clave y cada una de las dimensiones de calidad— en el sitio web de evaluación de IDRC, <https://www.idrc.ca/en/about-idrc/accountability/evaluation>. La infografía 'RQ+ de un vistazo' que aparece en la página anterior constituye un resumen de rápida consulta.

#### Las influencias clave

Este componente destaca aquellas influencias —ya sea dentro del proyecto de investigación o del entorno externo— que tienen mayores probabilidades de afectar la calidad de investigación. Estas influencias no pueden predecirse completamente si la evaluación es *ex ante*, pero esta sensibilización ante el contexto es uno de los aspectos más novedosos del Marco de Referencia RQ+. Considerar dichas influencias ayudará a que evaluadores, gestores, financiadores y otras partes intervinientes reflexionen de manera significativa y sistemática sobre los factores habilitantes o limitadores de la investigación y sobre el perfil de riesgo del proyecto, programa o portafolio. Así podrán, en lo posible, incorporar estas consideraciones en sus valoraciones.

El IDRC ha identificado cinco influencias clave sobre la calidad de investigación <sup>8</sup>.

## 1

**MADUREZ DEL CAMPO DE INVESTIGACIÓN**— El grado de existencia de marcos teóricos y conceptuales establecidos, a partir de los cuales se han formulado hipótesis bien definidas que han sido puestas a prueba. También se refiere a la existencia, y conjunto sustancial de investigaciones conceptuales y empíricas en su campo.

## 2

**FORTEALECIMIENTO DE LA CAPACIDAD PARA LA INVESTIGACIÓN**— La medida que indica hasta qué punto el proyecto de investigación se ocupa de fortalecer capacidades para la investigación por medio de apoyo financiero o técnico destinado a identificar y analizar desafíos de desarrollo y a concebir, realizar, gestionar y comunicar investigaciones que puedan abordar dichos desafíos.

## 3

**RIESGO EN EL ENTORNO INVESTIGATIVO**— Se refiere al contexto organizativo donde trabaja el equipo de investigación, y el grado en que dicho contexto respalda la investigación, donde el "respaldo" se relaciona a prioridades institucionales, incentivos e infraestructura.

## 4

**RIESGO EN EL ENTORNO POLÍTICO**— El grado de riesgos externos relacionados con factores potencialmente adversos que podrían surgir como resultado de problemas políticos o de gobernanza y que podrían afectar el proceso de investigación o el posicionamiento para su uso. Puede tratarse desde incertidumbre electoral e inestabilidad política hasta cuestiones más graves, como desestabilización política, conflictos violentos o crisis humanitarias.

<sup>8</sup> Se identificó una gran cantidad de influencias durante el proceso de desarrollo del IDRC, pero se las redujo a cinco. En otras organizaciones, estos factores pueden ser muy distintos.

## 5

**RIESGO EN EL ENTORNO DE DATOS**— Se indica la relativa disponibilidad y el consenso sobre la instrumentación y las medidas de recopilación y análisis de datos. Asimismo, se refiere al grado de existencia de datos (en abundancia o escasez de datos) en el entorno de investigación.

### Las dimensiones de la calidad de investigación.

El Marco de Referencia para la Evaluación RQ+ tiene cuatro dimensiones de calidad principales. Según nuestro punto de vista, estas dimensiones no están separadas, sino que existe una fuerte interrelación entre ellas. Hemos identificado y delimitado estas dimensiones a partir de nuestra intención de

#### RECUADRO 2 Ejemplos de influencias clave, con su escala

##### Madurez del campo de investigación

**(1) CAMPO ESTABLECIDO**

El campo se caracteriza por marcos teóricos y conceptuales afianzados y reconocidos; un grupo sustancial de investigaciones conceptuales y empíricas; medios de difusión distinguibles (revistas especializadas, conferencias, planes de estudio), y presencia de un cuerpo pujante de investigadores experimentados.

**(2) CAMPO EMERGENTE**

Reconocido por miembros y no miembros, con un cuerpo visible de trabajos, teorías y prácticas; medios de difusión distinguibles, y un modesto grupo de investigadores activos que se asocian fácilmente al campo y se conocen entre sí.

**(3) CAMPO NUEVO**

El campo de investigación tiene una base de conocimiento teórico o empírico limitada que aún está sujeta a debate o es muy cambiante; no tiene amplio reconocimiento; no cuenta con revistas especializadas ni programas académicos específicos, y solo tiene pocos investigadores activos que buscan reconocimiento.

##### Fortalecimiento de la capacidad de la investigación

**(1) ABORDAJE DÉBIL**

El fortalecimiento de la capacidad de la investigación no está presente o es de prioridad baja en este proyecto.

**(2) ABORDAJE MEDIO**

**(3) ABORDAJE SÓLIDO**

El fortalecimiento de la capacidad de la investigación es una de las prioridades importantes de este proyecto, junto con otras.

##### Riesgo en el entorno de datos

**(1) RIESGO BAJO**

Existe un amplio consenso y disponibilidad de instrumentación y medidas para la recopilación y el análisis de datos; el entorno de datos está bien desarrollado, es estable y hay abundancia de datos.

**(2) RIESGO MEDIO**

**(3) RIESGO ALTO**

La instrumentación y las medidas para la recopilación y el análisis de datos no se encuentran disponibles; las actividades de investigación se realizan en entornos muy inestables, poco avanzados y con escasez de datos.

##### Riesgo en el entorno investigativo

**(1) RIESGO BAJO**

El entorno investigativo —prioridades institucionales, incentivos, instalaciones, etc.— está afianzado y ofrece respaldo.

**(2) RIESGO MEDIO**

**(3) RIESGO ALTO**

El entorno investigativo es débil o está poco avanzado, y no ofrece respaldos.

##### Riesgo en el entorno político

**(1) RIESGO BAJO**

Entorno político estable con prácticas de gobernanza afianzadas, ausencia de conflictos destacables, etc.

**(2) RIESGO MEDIO**

**(3) RIESGO ALTO**

Entorno político muy inestable o volátil, con prácticas de gobernanza débiles, conflictos significativos, etc.

estudiar con mayor profundidad las cuestiones particulares que representan.

## 1

**INTEGRIDAD DE LA INVESTIGACIÓN—** Tiene en cuenta la calidad técnica, la adecuación y el rigor del diseño y la ejecución de la investigación según los métodos específicos y estándares comúnmente aceptados, y de acuerdo con lo reflejado en documentos del proyecto de investigación y en medios de difusión seleccionados. Las áreas destacadas incluyen el diseño de la investigación, el rigor metodológico, la revisión bibliográfica, el trabajo sistemático y la relación entre la evidencia recopilada y las conclusiones o las afirmaciones producidas. Los resultados revisados o no por pares son sometidos a diferentes procesos de valoración, empleando criterios diversos.

## 2

**LEGITIMIDAD DE LA INVESTIGACIÓN—** Tiene en cuenta hasta qué punto se han incluido las opiniones de los actores relevantes para producir los resultados de la investigación y si se consideró que el procedimiento fue justo en base en los valores, las inquietudes y las perspectivas de ese público. Se evalúa en relación a quiénes participaron y quiénes no; cómo fue el proceso de toma de decisiones; cómo se produjo, se aprobó y se difundió la información; la utilización de conocimiento local relevante, y si se respetaron las tradiciones y los sistemas de conocimiento locales. Se incluyen subdimensiones que instan al evaluador a tener en cuenta (i) las consecuencias potencialmente negativas para las poblaciones alcanzadas por la investigación, (ii) la respuesta ante cuestiones de género, (iii) la inclusión de las poblaciones vulnerables y (iv) la implicación del conocimiento local.

## 3

**IMPORTANCIA DE LA INVESTIGACIÓN—** Analiza la importancia y el valor de los conocimientos y entendimientos (aprendizajes) generados por la investigación en términos de la pertinencia percibida de sus procesos y productos para las necesidades y prioridades de los potenciales usuarios, así como el aporte de la investigación a la teoría o la práctica. Las subdimensiones se refieren a (i) la originalidad y (ii) la pertinencia de la investigación.

## 4

**POSICIONAMIENTO PARA EL USO—** Tiene en cuenta hasta qué punto o se ha dirigido el proceso de investigación y si los productos de la investigación se han preparado de manera tal que se ha mejorado su posibilidad de uso, influencia e impacto. La captación de una investigación es sobre todo un proceso político. Por lo tanto, para prepararla y favorecerla es necesario prestar atención al contexto de los usuarios, la accesibilidad de los productos y las estrategias de difusión e involucración adecuado al objetivo. También es necesario analizar detenidamente las relaciones que se deben establecer antes o durante el proceso de investigación, y las mejores plataformas para poner los resultados a disposición de determinados públicos y usuarios. El posicionamiento para el uso exige estrategias que integren a los potenciales usuarios en el proceso de investigación, siempre que esto sea factible y deseable. Las subdimensiones son (i) la accesibilidad e intercambio de conocimiento, y (ii) la oportunidad temporal (*'timing'*) y la usabilidad de los productos de la investigación.

Las primeras tres dimensiones —integridad, legitimidad e importancia de la investigación— son los factores fundamentales que se suelen encontrar, más o menos desarrollados, en la mayoría de los marcos de referencia para evaluar la calidad de investigación. La cuarta dimensión —el posicionamiento— es menos común y está representada por el 'plus' (RQ+) de este marco de referencia. Durante el proceso de creación del Marco de Referencia RQ+, IDRC y sus socios de investigación determinaron que sería razonable asignarse cierto nivel de responsabilidad para incrementar la probabilidad de uso de la investigación, es decir, por posicionar sus hallazgos para que ejerzan influencia y, eventualmente, contribuyan a generar impacto.<sup>10</sup>

### Las rúbricas de evaluación

La descripción de los niveles de las influencias clave (relativo al contexto) y el grado de desempeño de las dimensiones/subdimensiones de calidad se basan en rúbricas de evaluación adaptables que usan medidas cualitativas y cuantitativas. Para las influencias clave (véase el Recuadro 2), las rúbricas contienen texto descriptivo para guiar al evaluador para que pueda identificar características específicas que pueden encontrarse con diferente intensidad en un riesgo o factor determinado del contexto.

<sup>10</sup> La decisión de incluir esta dimensión se basó en el alcance de la misión y el mandato del IDRC, y también tuvo que ver con el período de programación de la investigación que el Marco de Referencia debía evaluar (un portafolio de proyectos que abarcaban entre tres y cinco años). Dependiendo de sus valores y su objetivo, otros financiadores u organizaciones de investigación pueden considerar que esta dimensión es poco realista, o tal vez deseen monitorizar la calidad con mayor énfasis en las esferas de influencia e interés.

Para las dimensiones de calidad y sus subdimensiones, la escala establece criterios de valoración que aclaran cómo se mide el desempeño. Al utilizar niveles graduados de éxito, la escala explica el significado de calificaciones como "bueno" o "inaceptable" en cada caso. El recuadro 2 incluye ejemplos de algunas de las escalas para dichas dimensiones.<sup>8</sup>

La síntesis de las calificaciones de las rúbricas en diferentes permutaciones permite clasificar el desempeño en cuatro niveles (desde "Inaceptable" hasta "Muy bueno") para conocer el estado de situación<sup>9</sup>, o para seguir el progreso con el paso del tiempo a medida que se fortalecen las capacidades de investigación y/o cuando los proyectos o programas están en camino a alcanzar la excelencia investigadora.

---

<sup>8</sup> Para acceder a la escala completa, véase: IDRC, 2014.

<sup>9</sup> Como se realizó en las revisiones externas de programas del IDRC en 2014/2015.

**RECUADRO 3** Ejemplos de las rúbricas para las dimensiones and subdimensiones de calidad

**Dimensión 1.0: Integridad de la investigación**

INACEPTABLE		MENOS QUE ACEPTABLE		DE ACEPTABLE A BUENO		MUY BUENO	
1	2	3	4	5	6	7	8
La investigación tiene poco o ningún mérito científico. La justificación del enfoque utilizado es cuestionable. Existen brechas graves en el rigor metodológico del análisis bibliográfico, así como en la recopilación y el análisis de datos.		Se evidencian intentos de cumplir con los estándares metodológicos, pero no del todo exitosos. Existen deficiencias importantes al justificar la elección del diseño y los métodos de investigación.		Se cumplen los estándares metodológicos para el diseño y la ejecución de la investigación.		El mérito científico es incuestionable. Existe evidencia de una exhaustividad excepcional en el diseño de la investigación y en todas sus fases de ejecución. El producto es un ejemplo de lo que significa cumplir estos criterios.	

**Dimensión 2: Legitimidad de la investigación; Subdimensión 2.4: Implicación del conocimiento local**

NO CORRESPONDE	INACEPTABLE		MENOS QUE ACEPTABLE		DE ACEPTABLE A BUENO		MUY BUENO	
	1	2	3	4	5	6	7	8
Debido a la naturaleza de la investigación, no es necesario tener en cuenta o implicar el conocimiento local.	Se ha descuidado la implicación con los contextos locales durante el proceso de investigación. Se pueden encontrar varios puntos débiles importantes en relación con (i) la forma en que se identificaron las necesidades y las preguntas de investigación, (ii) cómo se involucró a las comunidades o poblaciones locales, (iii) cómo se consideraron los contextos y los sistemas de conocimientos locales, y (iv) cómo se aseguraron los beneficios locales del proceso de investigación.		Se han tenido en cuenta los contextos locales durante el proceso de investigación, pero persisten algunos puntos débiles en relación con (i) la forma en que se identificaron las necesidades y las preguntas de investigación, (ii) cómo se involucró a las comunidades o poblaciones locales, (iii) cómo se consideraron los contextos y los sistemas de conocimientos locales, y (iv) cómo se aseguraron los beneficios locales del proceso de investigación.		El contexto y la implicación local han sido parte de los enfoques del proceso de investigación. Quedan muy pocos, si acaso, puntos débiles menores en relación con (i) la forma en que se identificaron las necesidades y las preguntas de investigación, (ii) cómo se involucró a las comunidades o poblaciones locales, (iii) cómo se consideraron los contextos y los sistemas de conocimientos locales, y (iv) cómo se aseguraron los beneficios locales del proceso de investigación.		El contexto y la implicación local han sido centrales en el proceso de investigación. Se identificaron adecuadamente las necesidades y las preguntas de investigación. Se involucró a las comunidades o poblaciones locales. Se consideraron y respetaron los contextos y los sistemas de conocimientos locales. Y se aseguraron los beneficios locales del proceso de investigación.	

**Dimensión 3: Importancia de la investigación; Subdimensión 3.2: Pertinencia**

INACEPTABLE		MENOS QUE ACEPTABLE		DE ACEPTABLE A BUENO		MUY BUENO	
1	2	3	4	5	6	7	8
Existe poca o ninguna evidencia de que la investigación pudiese contribuir a una prioridad local, a una política o estrategia de desarrollo clave, o a un área emergente que pueda requerir soluciones en un futuro previsible. Las evaluaciones de necesidades y la justificación del trabajo no existen o son poco convincentes.		Existe alguna evidencia de que la investigación pudiese contribuir a una prioridad local, a una política de desarrollo clave o a un área emergente que pueda requerir soluciones en un futuro previsible. El tratamiento de esta área de trabajo parece tener una justificación razonable en la actualidad.		Existe evidencia sólida de que la investigación pudiese contribuir a una prioridad local, a una política o estrategia de desarrollo clave, o a un área emergente de relevancia que pueda requerir soluciones en un futuro cercano. El tratamiento de esta área de trabajo, parece tener muy buena justificación en la actualidad.		Existe evidencia sólida de que ya se reconoce que la investigación tiene el potencial de abordar una prioridad local crítica, una política o estrategia de desarrollo clave, o un área emergente de importancia que tiene muchas probabilidades de requerir soluciones en un futuro cercano. El tratamiento de esta área de trabajo en la actualidad coloca a los investigadores a la vanguardia de un campo de trabajo activo o significativo.	

**Dimensión 4: Posicionamiento para el uso; Subdimensión 4.2: Oportunidad temporal (*timing*) y usabilidad**

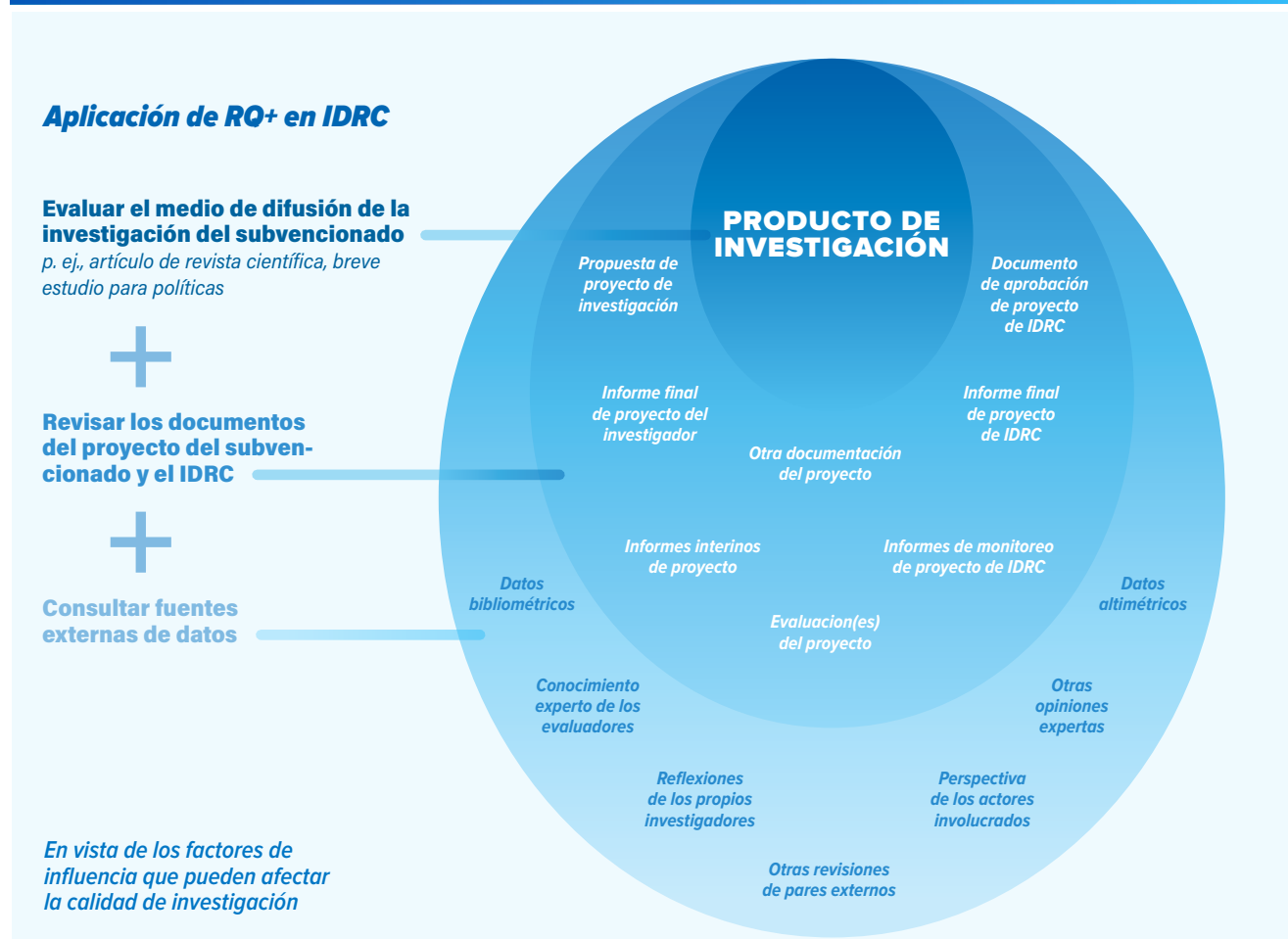
INACEPTABLE		MENOS QUE ACEPTABLE		DE ACEPTABLE A BUENO		MUY BUENO	
1	2	3	4	5	6	7	8
Existe poca o ninguna evidencia de que se haya realizado algún análisis del entorno de usuarios relevantes y de que se hayan considerado las contingencias organizativas, políticas, sociales o económicas.		Existe evidencia de que se realizaron algún análisis del entorno de los usuarios. Sin embargo, no se los ha tenido en cuenta en su totalidad y, además, los análisis no están acompañado de estrategias o planes para transferir el conocimiento generado hacia políticas o prácticas.		Existe evidencia de que se han analizado el entorno de los usuarios y las contingencias principales, se ha reflexionado acerca de estos factores. Esto se ha conectado estrategias o planes para transferir oportunamente el conocimiento generado hacia políticas o prácticas.		El análisis del entorno de los usuarios y de las contingencias es exhaustivo y excepcionalmente bien documentado y articulado. Existe evidencia de que se han evaluado detenidamente las posibilidades de éxito de estrategias diseñadas para abordar contingencias.	

## Utilización del Marco de Referencia

Los productos de investigación son la unidad de análisis de muchos marcos de evaluación de investigaciones que dependen fuertemente de la revisión por pares. **La premisa del Marco de Referencia para la Evaluación RQ+ es que, para producir una evaluación creíble, equilibrada e integral de la calidad de investigación, se deben considerar elementos que trascienden los productos de la investigación e incluir aspectos clave del contexto donde ésta se lleva a cabo, junto con la manera en que se ha gestionado.** En consecuencia, la aplicación del presente Marco de Referencia combina el análisis contextual con la valoración del producto y del propio proyecto de investigación.

Como en la mayoría de los procesos de evaluación de la investigación, la revisión por pares expertos juega un papel importante cuando se aplica el Marco de Referencia. Expertos externos de campos relevantes de investigación deben juzgar, por ejemplo, el grado de cumplimiento de cada subdimensión. También deben recopilar datos y evidencia para describir mejor el contexto de la investigación, y para verificar y profundizar sus hallazgos basándose en múltiples métodos y fuentes. Durante la aplicación del Marco de Referencia por parte de IDRC en las Revisiones Externas sumativas de 2014/2015, se analizaron los archivos de los proyectos y los programas, y se realizaron encuestas y entrevistas entre los informantes clave (véase la Figura 3).<sup>11</sup>

**FIGURA 3** Recopilación de datos necesarios para implementar el RQ+ en IDRC



<sup>11</sup> Los informantes destacados pueden ser expertos en el campo, con conocimiento del contexto donde se realiza el trabajo, los usuarios previstos, los actores principales (como el personal de los proyectos y los programas) e, idealmente, cuando sea factible y esté justificado, los participantes en la investigación.



## Un enfoque escalonado para implementar RQ+

Utilizando un instrumento de valoración personalizado, IDRC implementó el Marco de Referencia RQ+ en cuatro pasos principales, tal como se muestra en la Figura 4. Estos se describen brevemente a continuación<sup>12</sup>.



### PASO 1

#### *Seleccionar la muestra*

La estrategia de muestreo para el portafolio de un programa tiene dos niveles: i) seleccionar una muestra de proyectos de investigación del portafolio que cumplan ciertos criterios y, entonces, ii) seleccionar un conjunto de productos de investigación que serán el foco de la evaluación. En ambos casos, pueden tomarse las muestras deliberadamente o al azar. Para que una muestra sea representativa del portafolio de un programa, los revisores deben agrupar y mapear los proyectos de acuerdo en base a atributos seleccionados, como la temática, monto presupuestario, período (tiempo) de implementación, tipo de productos, público o interés suscitados (potenciales), el lugar de la investigación, el lugar de los efectos previstos o logrados, y el tipo de efectos. Se sigue el mismo tipo de proceso al seleccionar los productos.

### PASO 2

#### *Caracterizar los proyectos seleccionados con base en las influencias clave*

Durante el Paso 2, los revisores utilizan la documentación del proyecto, su propia experiencia y conversaciones con informantes clave para caracterizar cada uno de los proyectos seleccionados según las influencias clave que tengan más probabilidades de afectar la calidad de investigación. Los revisores describen cada influencia utilizando una rúbrica (Recuadro 2). Pueden utilizarse diagramas de grupos (o *clústeres*) o ayudas visuales similares para mapear y comparar las influencias dentro de proyectos y portafolios de programas, y entre unos y otros. También sirven para analizar su influencia sobre el progreso esperado hacia la excelencia en materia de investigación. Esta manera de visualizar los datos también permite que los evaluadores y los directores de programas mapeen sistemáticamente los niveles y los tipos de riesgo, y consideren estrategias para encauzar la gestión y la evaluación de los portafolios.

<sup>12</sup> Se puede acceder a información más detallada en IDRC, 2014.

## PASO 3

### *Calificar la calidad de investigación*

El Paso 3 del proceso de evaluación consiste en calificar la calidad de los medios de difusión de la muestra del proyecto, utilizando las dimensiones y subdimensiones de calidad incluidas en el Marco de Referencia. Cada subdimensión se puntúa dentro de una escala de ocho puntos (ejemplos en el Recuadro 3). Las rúbricas se diseñaron teniendo en cuenta los valores del IDRC y su interpretación institucional de cada dimensión y subdimensión. La rúbrica clarifica los criterios y las normas para la calificación a la vez que establece la base de evidencia, para así sintetizarlos en criterios generales de evaluación.

Los revisores deben recibir pautas o tomar decisiones sobre el nivel de coherencia y estandarización de los criterios, y sobre las diferentes ponderaciones que deben asignar a las distintas dimensiones (y subdimensiones) para los proyectos y los programas. Esto es esencial y, como se expone a continuación, estas decisiones pueden suponer un desafío.

## PASO 4

### *Sintetizar las calificaciones*

En el último paso del proceso de evaluación, los revisores agregan las calificaciones de cada subdimensión para obtener el total por dimensión para cada producto. Luego se agregan los resultados, primero a nivel del proyecto y, después, a nivel del portafolio del programa. Con las fórmulas correspondientes, pueden también agregarse las calificaciones obtenidas de un conjunto de proyectos hasta llegar al nivel de programa, por dimensión, subdimensión o por factores de influencia. De esta manera, se pueden destacar las fortalezas y los puntos débiles en ciertas dimensiones de la calidad de investigación, para tipos específicos de proyectos o en contextos puntuales (según lo identificado al analizar los factores de influencia). De hecho, es posible elevar el nivel de agregación a varios programas o inclusive en el nivel de la organización como un todo.<sup>13</sup>

Las calificaciones totales del portafolio de proyectos de un programa pueden registrarse en una tabla. La escala de ocho puntos permite clasificar el desempeño en cada dimensión; las calificaciones van desde el nivel más bajo ('inaceptable') hasta el más alto ('muy bueno'). Si las calificaciones cuantitativas resultan demasiado chocantes, se las puede sustituir por un código de colores ('semáforo'). De hecho, la síntesis de datos cualitativos puede ser igual de importante —o incluso más— que la síntesis de calificaciones cuantitativas, dependiendo de los objetivos de aprendizaje de la evaluación. Si estos datos cualitativos son más relevantes que los otros y forman la base de las calificaciones, estos son los principales datos que se deben sintetizar y analizar. En las síntesis cualitativas o cuantitativas, las compilaciones pueden considerarse como un valor total combinado. (En la síntesis cuantitativa, este valor no necesariamente es la media, la mediana o la moda). La ponderación del valor total depende del contexto y de los objetivos que tengan los proyectos de investigación, el portafolio del programa o la organización.

A efectos de la gestión, se puede utilizar las agregaciones de las influencias clave y las dimensiones/subdimensiones de la calidad de investigación para identificar patrones o posibles correlaciones entre los datos. Por ejemplo, un grupo de proyectos identificados como 'de alto riesgo' según el análisis de factores de influencia puede compararse con las calificaciones de calidad de investigación para determinar si algún factor específico, o combinación de factores, influyeron sobre dimensiones específicas de la calidad.

<sup>13</sup> El IDRC aún no ha sometido a prueba estos niveles de compilación.



## Algunas Reflexiones Sobre Pros y Contras en el Diseño del Marco de Referencia

Al diseñar cualquier marco de referencia para evaluar la investigación, es inevitable realizar sacrificios y definir prioridades. No existen soluciones mágicas. Cuando se desarrolló el Marco de Referencia RQ+, se tomaron varias decisiones en base a comentarios recibidos por parte de los actores interesados durante las fases iniciales de diseño y pruebas.

Reconocimos que el carácter integral del Marco de Referencia y el trabajo sistemático y detallado necesario durante el proceso de evaluación tienen costos relativos a la facilidad de uso y los tiempos necesarios. No obstante, si lo comparamos con la revisión por pares convencional, el enfoque necesario para juzgar las dimensiones de la calidad de investigación permite realizar valoraciones más claras, ricas y matizadas, y, por tanto, es decir, más precisas y transparentes. Esto se logra, principalmente, utilizando una escala y múltiples fuentes de evidencia que exigen la integración e interpretación sistemática de datos cualitativos y cuantitativos.

También tuvimos en cuenta la flexibilidad del marco de referencia, ya que puede ser personalizado con las subdimensiones y rúbricas, de acuerdo con los motivos explicados en los fundamentos de su diseño. Sin embargo, esto puede complicar las aplicaciones cuando se requiera una normalización estricta de las mediciones o cuando diferentes organizaciones e iniciativas pretendan agregar y comparar las calificaciones.

Un tercer aspecto se refiere a la flexibilidad de asignar diferentes ponderaciones a cada una de las dimensiones y subdimensiones (y a las influencias clave) de la calidad, en lugar de peticiones recibidas para estandarizarlas entre distintas organizaciones. El alcance de la ponderación depende en gran medida de los valores de la organización que aplique el Marco de Referencia. Por ejemplo, durante la fase de diseño en el IDRC, algunos empleados y revisores del programa que testearon el Marco de Referencia consultaron si la dimensión 'Integridad de la investigación' debía tener una ponderación mayor, ya que temían que, de lo contrario, el mérito académico clásico se vería devaluado. Dada la política corporativa del IDRC sobre las publicaciones de acceso abierto, también se discute si deben asignarse calificaciones más altas a la dimensión de "Posicionamiento para el uso" en aquellos proyectos y programas que ofrezcan este tipo de publica-

ciones. Sin embargo, esto puede causar problemas con otros programas y provocar desacuerdos con socios financiadores que consideren que la publicación de la investigación en prestigiosas revistas científicas es el principal indicador de calidad, incluso si esas revistas no tienen políticas de acceso abierto.

Los revisores que participaron en la primera aplicación del Marco de Referencia para la Evaluación RQ+ confirmaron las ventajas de varias de estas decisiones, pero también pusieron de manifiesto las complicaciones que creaban estas compensaciones al realizar la evaluación. La siguiente sección aborda las lecciones y experiencia obtenida hasta la fecha en la aplicación del Marco de Referencia en IDRC.

## PARTE 3

# Lecciones de la Implementación del Marco de Referencia RQ+ por IDRC

### El Contexto de la Primera Aplicación

La evaluación externa es uno de los pilares del sistema de evaluación del IDRC. Las Revisiones Externas de programas son uno de los procesos de mayor nivel de rendición de cuentas que tiene el Centro. Al momento de esta publicación, estas se estructuraban como evaluaciones sumativas de cada programa del IDRC y eran llevadas a cabo cada cinco años por un equipo de tres expertos en el tema, independientes y externos, dirigidos por la División de Políticas y Evaluación de IDRC. Los equipos de revisores estaban conformados por experimentados investigadores y evaluadores con amplia experiencia académica en la revisión por pares. Las Revisiones tienen dos propósitos principales. En primer lugar, respaldan la rendición de cuentas del IDRC en la ejecución de programas, y, en consecuencia, la Junta de Gobernadores de IDRC es su principal usuario. En segundo lugar, aportan nuevos conocimientos a los directores de programas, para asistirles en la gestión de sus portafolios de subvenciones.

Las revisiones externas son evaluaciones sumativas e integrales que analizan la implementación, el desempeño y los efectos de los programas. Abordan cuatro elementos: i) la coherencia, efectividad y adecuación de las estrategias de implementación de programas; ii) *la calidad de investigación apoyada por cada programa revisado*; iii) la pertinencia e importancia de los efectos de los programas, y iv) los asuntos que deben ser considerados por la Junta de Gobernadores y el equipo de dirección del IDRC. Como la calidad de investigación era uno de estos cuatro elementos, las Revisiones Externas presentaban una oportunidad única de aplicar el RQ+ y lograr un grado de estandarización para los criterios empleados en las evaluaciones de programas.

Antes de ser aplicado en las Revisiones, se puso a prueba el Marco de Referencia mediante una serie de simulaciones basadas en programas que ya se habían revisado en 2009 y 2010. Las simulaciones apuntaban a identificar debilidades significativas antes de que pudieran afectar el conjunto de Revisiones que se estaba por lanzar. Durante un período de tres meses, a principios de 2014, cuatro antiguos revisores externos que conocían IDRC, el proceso de Revisión Externa y los programas seleccionados<sup>14</sup> aplicaron el Marco de Referencia y documentaron sus experiencias sobre una amplia gama de aspectos. Entre ellos estuvieron el grado en el que dicho Marco permitía una valoración justa, su idoneidad para distintos tipos y fines de investigación, y la viabilidad de su uso para evaluaciones de programas dentro de los parámetros y los plazos de un proceso de Revisión Externa. Las experiencias de simulación se compartieron en un taller con un grupo de directores de programa intermedios y *senior* de IDRC. Con base en las lecciones aprendidas, se mejoró y se aplicó el Marco de Referencia en siete Revisiones Externas del IDRC.

En el resto de este documento, se presenta la experiencia de IDRC durante la última ronda de Revisiones Externas, realizadas entre junio de 2014 y septiembre de 2015. Las siguientes subsecciones describen lo aprendido en relación a logros, desafíos y posibles rumbos futuros.

<sup>14</sup> Uno de los revisores también fue uno de los diseñadores del Marco de Referencia RQ+.

## Motivos para Ser Optimistas

### *Los valores incorporados se convierten en explícitos*

El RQ+ fue fruto de un proceso de reflexión interno de IDRC. Este proceso se orientó a producir criterios de evaluación que pudieran encapsular los valores y los conceptos compartidos del Centro sobre calidad de investigación. Así, el propio Marco de Referencia orientó a los revisores externos hacia temas que eran pertinentes, valorados y específicos para IDRC. El proceso de revisión, al utilizar un instrumento que captaba la esencia de los valores de IDRC, produjo evaluaciones apreciadas por los actores primarios involucrados como muy útiles. Los revisores también se beneficiaron. Se sintieron cómodos al emitir juicios en base a criterios que los usuarios principales identificaron como importantes y que los beneficiarios de subvenciones del IDRC debían cumplir durante su investigación.

Durante sus presentaciones finales, los revisores manifestaron que era posible transferir este aspecto del enfoque RQ+ para que fuera aplicado por otras entidades de investigación. La estructura general y el proceso de implementación del enfoque pueden estar estandarizados, pero las subdimensiones y los factores de influencia pueden cambiarse en mayor o menor medida para reflejar los valores o las prioridades de aprendizaje de otra organización. La implementación del Marco de Referencia también puede tener beneficios secundarios, como establecer un sistema de valores común en la organización o entre distintas organizaciones, y promover las metas deseadas en la gestión y administración de las investigaciones.

### *La naturaleza sistemática del enfoque consolidó los procesos y los resultados de la evaluación*

En este punto, observamos optimismo con respecto a los procesos y a los productos. En cuanto al proceso, el RQ+ proporcionó instrucciones detalladas y claras a los revisores. La amplia utilización de rúbricas arrojó luz sobre la base de las evaluaciones y aportó mayor coherencia a los criterios de la que suele existir en procesos convencionales de revisión por pares. Los revisores consideraron que esta era una ventaja valiosa, a pesar del considerable nivel de análisis necesario para utilizar las rúbricas en cada subdimensión y factor de influencia. Muchos revisores indicaron que este 'proceso guiado' simplificó la evaluación.

También se percibió que el Marco de Referencia afianzaba el resultado de la Revisión Externa. Se reconoció que ayuda a abordar (aunque no en su totalidad) una dificultad específica de las evaluaciones de investigación: la posibilidad de comparar tipos muy diferentes de investigaciones. Como expresó uno de los revisores, el Marco de Referencia proporcionó *"cierta comodidad al comparar los resultados de una prueba controlada aleatorizada con los resultados de un proyecto de investigación impulsado por el cabildeo (advocacy) o la acción participativa"*. Para los revisores esto constituyó un aporte útil y novedoso. Esto se debe a la integración de varias fuentes de datos, un conjunto estable de dimensiones y subdimensiones para evaluar la investigación, y el reconocimiento explícito de la necesidad de incorporar factores contextuales al análisis.

Los revisores también valoraron que el proceso sistemático y detallado de aplicación del Marco de Referencia no solo aumentó su confianza en la calidad de Revisión, sino que mejoró las probabilidades de replicabilidad. Esta observación es ilusionante si se tienen en cuenta los debates y las dudas actuales en torno a la replicabilidad de las revisiones por pares.

### *Trascender el producto de investigación permite mayor precisión, exactitud, y grado de detalle*

La experiencia de las Revisiones Externas demostró que el RQ+ facilita análisis precisos y detallados, gracias a la naturaleza integral del enfoque. Esta ventaja se torna evidente en lo que los revisores denominaron 'trascender el producto de investigación'. Aunque es más complejo, representa una mejora sobre los sistemas tradicionales de revisión por pares ya que estos últimos están restringidos al objeto en revisión. Por ejemplo, un documento de investigación presentado ante una revista académica es evaluado solo según el mérito del documento en sí. De la misma manera, al postularse a una subvención para la investigación, es probable que solo se evalúe una propuesta de investigación y, tal vez, los antecedentes del postulante o su equipo. En estos dos ejemplos, a los que llevan a cabo la revisión por pares se les indica que recurran a su experiencia y a los conocimientos adquiridos, y, en muchos casos, que no tengan en cuenta nada más. Parece irónico que la evaluación de una investigación pueda realizarse con tan poca observación empírica, es decir, recopilación, validación y análisis de datos.<sup>15</sup>

<sup>15</sup> Se aclara que la experiencia de los revisores también se podría clasificar en dimensiones de RQ+, por ejemplo, como investigadores especialistas, usuarios del conocimiento o investigadores de disciplinas complementarias. Las revisiones externas del IDRC que se describen aquí no incluyeron este paso.

Para resolver este problema, la implementación del RQ+ requiere esta evaluación por pares, pero también se deben tomar los datos del entorno donde se ha realizado la investigación. Las entrevistas cualitativas con investigadores que realizaron el proyecto, directores de investigación y usuarios actuales o potenciales (formuladores de políticas y profesionales)<sup>16</sup> aportaron detalles y mayor exactitud a la evaluación tanto a nivel de proyectos como de programas. Otros métodos notables utilizados por los revisores para ‘ir más allá del producto’ fueron escaneos bibliométricos y altimétricos, así como encuestas personalizadas a actores influyentes en el propio campo de investigación.

## Desafíos e Inconvenientes

La aplicación del RQ+ en IDRC también puso en evidencia ciertos desafíos notorios.

### *La evaluación insume mucho tiempo, en especial cuando se debe realizar una triangulación sólida*

La implementación del enfoque RQ+ requiere sintetizar datos cualitativos y cuantitativos de múltiples fuentes y métodos. Se demostró que las consultas amplias con actores internos y externos son esenciales para cubrir brechas y permitir una triangulación suficiente. Por lo tanto, la calidad y la accesibilidad de las fuentes de información de los proyectos y los programas son factores críticos para la viabilidad y el valor del enfoque. Además, cuando se establecen sistemas de monitoreo sin hacer referencia a componentes del Marco de Referencia, la recopilación de datos puede insumir una gran cantidad de tiempo y recursos. Los revisores consideraron que si se tiene en cuenta el equilibrio entre el carácter integral, la facilidad de uso y los resultados confiables, frente al tiempo adicional necesario, el balance es positivo. No obstante, al aplicar el Marco de Referencia, debe planificarse a conciencia el tiempo y el esfuerzo que destinarán.

### *Es crítico que el muestreo sea riguroso y creíble*

Las Revisiones Externas de IDRC requieren la selección de una muestra de proyectos de investigación del portafolio de cada programa. La estrategia de muestreo es crucial para la credibilidad de la iniciativa y puede necesitarse un alto nivel de concienciación para ejecutarla. Por lo tanto, debe

llevarse a cabo con atención y esmero para garantizar una representatividad justa del programa de investigación siendo evaluado.

El muestreo en la aplicación del enfoque en IDRC se realizó mediante un enfoque deliberado, implementado de forma iterativa. Se compartieron y debatieron los criterios para seleccionar los proyectos de investigación con los directores de los programas antes de tomar una muestra final. Aunque este puede no ser el mejor abordaje bajo todas las circunstancias, funcionó bien en este contexto. Muestras aleatorias pueden ser perfectamente adecuadas y preferibles para seleccionar proyectos en otros escenarios.

Tomar muestras de portafolios de programas complejos y de gran envergadura puede presentar dificultades técnicas y políticas, y, para obtener los mejores resultados, los interesados y los usuarios principales deben considerarlas legítimas y creíbles. Para optimizar los debates y controlar las expectativas, es necesario que, desde el principio, se seleccione una muestra, de forma deliberada o al azar, que sea creíble y aceptada por todos los interesados principales, o que sea aprobada por la autoridad definitiva. En el contexto del IDRC, fue muy valioso realizar consultas con el personal de los programas. Por supuesto, debe procurarse mitigar posibles sesgos. Si el objetivo de la evaluación exige un proceso de toma de muestras independiente del personal del programa, otros actores interesados pueden participar a definir la muestra.

### *La cuantificación, después de combinar datos cuantitativos y cualitativos, puede arrojar resultados de apariencia simplista*

A nivel micro, el enfoque RQ+ establece que los revisores evalúen los procesos de investigación utilizando datos cualitativos y cuantitativos. Se consideró que las rúbricas eran útiles para producir criterios más precisos, incluyendo para combinar los dos tipos de evidencia (*cuali/cuanti*). No obstante, este proceso se tornó problemático cuando los resultados se expresaron en valores numéricos (las calificaciones de cada rúbrica). Puede decirse que si no había una referencia al texto específico de las rúbricas, se percibía que las calificaciones numéricas no captaban adecuadamente el rigor y profundidad de los resultados, y, por consiguiente, del auténtico valor y espíritu de la evaluación. Algunos revisores intentaron mitigar esta percepción utilizando un código de colores en vez de calificaciones cuantitativas. Las dificultades aumentaban

<sup>16</sup> No se incluyó a los participantes de la investigación en las revisiones externas, pero, de existir el tiempo y los recursos necesarios, sería valioso incorporarlos, en línea con el espíritu del Marco de Referencia.

cuando algunas subdimensiones eran 'no aplicable' en ciertos proyectos o productos.

A nivel macro, la comparación y la agregación de datos presentaba dos complicaciones: i) comprender los valores relativos de las calificaciones entre las subdimensiones y decidir cómo se los debía ponderar y valorar, y ii) manejar la incertidumbre, ya que el Marco de Referencia orienta la síntesis de las evaluaciones de proyectos hacia una evaluación a nivel de programas, partiendo de calificaciones numéricas basadas en rúbricas. El valor de una rúbrica al establecer un promedio para todo el programa o una valoración integral para los factores de influencia o las subdimensiones en un programa se podían interpretar, y se interpretaron, tanto en términos positivos como negativos por parte de los revisores y del personal de los programas.

Por el lado positivo, las calificaciones de las rúbricas representaban de manera bien definida y directa la calidad de investigación para las diferentes subdimensiones de interés. Para un usuario de la evaluación interesado en la rendición de cuentas y los resultados sumativos, este hallazgo puede resultar valioso y útil. Por el lado negativo, se consideró que las rúbricas simplificaban demasiado la complejidad de la evaluación. Para el director de un programa de investigación que deseaba datos sobre *por qué o cómo* se podría mejorar la calidad de investigación, los datos detrás de la calificación de la rúbrica podrían resultar más interesantes.

Vale la pena mencionar que los revisores y los directores de programas sometidos a revisión articularon estas tensiones con respecto a las rúbricas.

---

*Utilizar el Marco de Referencia para la Evaluación RQ+ en comparaciones entre programas puede tener consecuencias imprevistas y no deseadas.*

---

Del punto anterior se desprende la utilización de la agregación de rúbricas como un mecanismo para comparar programas. Por ejemplo, si el usuario final de la evaluación es el equipo directivo o el consejo de dirección de una organización, una comparación de estas características sería útil en una valoración estratégica e integral del desempeño de la investigación. Esto puede resultar especialmente útil si el debate más amplio sobre el desempeño es respaldado por la

combinación de datos cualitativos y cuantitativos que ofrecen las rúbricas. No obstante, existe el riesgo de que las comparaciones basadas en una interpretación simplista de calificaciones numéricas puedan desencadenar una competencia entre programas y, sin querer, abra paso a una "olimpíada de programas". Esto podría producir ansiedades no deseadas en el personal de los programas y hasta una caída en picado de las mediciones de desempeño.<sup>17</sup> Esta ansiedad es contraria a la creencia fundamental del IDRC de que el monitoreo y la evaluación ofrecen métodos útiles y constructivos de valoración interna y externa que permiten aprender, rendir cuentas, mejorar y lograr resultados deseados.

---

<sup>17</sup> Este riesgo existe, pero no se dio un caso tan drástico en el proceso de revisión externa del IDRC.



## PARTE 4

# El Potencial del Marco de Referencia RQ+

### Los Múltiples Propósitos del Marco de Referencia

La primera aplicación del RQ+ en las revisiones externas del IDRC demostró que puede aplicarse con éxito para evaluaciones sumativas donde la calidad de investigación sea una prioridad. Sin embargo, esta experiencia también permitió descubrir aplicaciones alternativas del Marco de Referencia en el proceso de investigación.

---

*Puede facilitar la creación de alianzas y la planificación coordinada entre socios financiadores o socios de investigación*

---

Al definir las prioridades, valores y principios sobre los cuales se juzga la calidad de investigación, el Marco de Referencia se puede utilizar para sentar bases comunes entre los asociados que financien o evalúen iniciativas juntas. Esto se ejemplifica en la dimensión "Legitimidad de la investigación" y sus rúbricas. IDRC financia investigaciones para el desarrollo en el Sur Global. Por consiguiente, la aplicación del Marco de Referencia, en este caso, reconoció expresamente la importancia de hacer hincapié en el conocimiento local, la inclusión y la sensibilidad ante cuestiones de género, junto con la necesidad de prestar atención a las posibles consecuencias negativas de la investigación. Los socios pueden tener los mismos valores y prioridades subyacentes o no. Por ello, los debates sobre cómo se usará el Marco de Referencia —y quiénes lo utilizarán— pueden ofrecer oportunidades para decidir cómo acomodar (o no) las diferencias. Por ejemplo, la dimensión "Integridad de la investigación" incorpora consideraciones del mérito científico tradicional que suelen ocupar un lugar importante en la manera en que se concibe la calidad de investigación. Durante la primera aplicación del Marco de Referencia, algunos revisores se preguntaron si la integridad de la investigación debía extenderse más o si

debía recibir una mayor ponderación.<sup>18</sup> También debatieron la posibilidad de incorporar una dimensión de eficiencia de la investigación sobre el nivel de aprovechamiento de los recursos, al Marco de Referencia.<sup>19</sup> Todas estas reflexiones son legítimas y pueden conformar una plataforma para cualquier debate sobre 'lo que importa' en el desempeño y la evaluación de las investigaciones.

Como ya se mencionó, una de las ventajas del Marco de Referencia es su flexibilidad. Puede adaptarse a distintos contextos de referencia y necesidades de gestión. No todas las dimensiones de la calidad de investigación deben evaluarse en todas las situaciones y pueden crearse nuevas dimensiones que reflejen los mandatos específicos de una organización. De cualquier manera, el Marco puede estandarizarse, lo que facilita la consistencia y la comparación entre proyectos, programas y organizaciones.

---

*Puede servir de base para decisiones de financiación y consolidar proyectos de investigación en curso*

---

El RQ+ puede contribuir a la elaboración y evaluación de propuestas de investigación presentadas en respuesta a concursos o convocatorias. Los oficiales de programas y los encargados de realizar revisiones por pares también pueden usar el Marco de Referencia para llevar a cabo valoraciones sistemáticas y sostenidas de criterios deseados al principio o en el transcurso de un proyecto o programa. Los directores de proyectos de investigación lo pueden utilizar para garantizar que los protocolos de investigación en su ámbito incluyan dimensiones importantes para ellos y, además, para asegurarse de que el equipo de investigación haya tenido en cuenta factores que pueden influir en los buenos resultados y que hayan tomado medidas para garantizar la traducción final de los conocimientos.

---

<sup>18</sup> En caso de que se apliquen ponderaciones a diferentes dimensiones del marco de referencia. Para sus revisiones externas, el IDRC decidió no asignar ponderaciones.

<sup>19</sup> El IDRC reconoce que esta dimensión podría agregar valor real al marco de referencia, pero no estaba conforme con la métrica que se utiliza actualmente y, además, no le pareció que se adaptara a su mandato o a su modelo de negocio.

Además, si el Marco de Referencia se ha utilizado en simultáneo para conformar una base de datos de monitoreo de la calidad de investigación, pueden tenerse en cuenta las fortalezas y debilidades de un programa de investigación al decidir qué financiar o cómo estructurar protocolos de investigación propuestos.

---

#### *Puede facilitar la meta-evaluación y el meta-análisis*

---

El enfoque sistemático y detallado para recopilar y analizar datos que exige el enfoque RQ+ también puede facilitar la síntesis y la meta-evaluación. Por ejemplo, los resultados de investigación entre programas pueden combinarse para obtener una perspectiva organizativa (corporativa) sobre la calidad de investigación o sobre dimensiones específicas de su desempeño. Poco después de la publicación de este documento, a principios de 2016, el IDRC se encontrará en las fases de planificación de este tipo de meta-evaluación al examinar sus Revisiones Externas. Por poner otro ejemplo, a una persona que aporta fondos para una investigación y que utiliza el Marco de Referencia para evaluar nuevas postulaciones le puede resultar útil monitorizar la situación (agregando las valoraciones de cada postulación) de las áreas en las cuales la comunidad demuestra fortalezas y debilidades, en especial si la organización ha priorizado el fortalecimiento de la capacidad investigativa. Una revista académica que utilice una adaptación del Marco de Referencia para estructurar la revisión por pares de sus manuscritos puede combinar las valoraciones anuales para comprender mejor los defectos y las virtudes del cuerpo de investigadores o de la disciplina científica que busca promover.

Por último, los enfoques de la evaluación de investigaciones que estandarizan los criterios para una organización, un programa, una revista académica o un campo de estudio — como puede hacerlo el Marco de Referencia RQ+— tienen el potencial de facilitar el meta-aprendizaje y la síntesis de conocimiento al analizar temas y cuestiones de interés transversal. Por ejemplo, un tema de interés como las cuestiones de género, puede compilarse y compararse entre distintas disciplinas o programas de investigación, o incluso entre distintos años.

---

#### *Puede ayudar a comprender mejor cómo se logra el impacto de la investigación*

---

La naturaleza integral y sistemática del enfoque RQ+ tiene el potencial de recabar información que puede ser útil al estudiar el impacto de las investigaciones, que suelen financiarse con la intención de producir un impacto de largo plazo para el bien común. Asimismo, la investigación puede producir beneficios o costos potenciales que pueden visibilizarse mucho después de concluido el proyecto. Aunque el Marco de Referencia se centra solo en la esfera de control de los actores primarios, la información recopilada durante la valoración constituye una imagen o visualización amplia del estado del proyecto, programa u organización en un momento dado. Al combinar en retrospectiva esta información con el trazado de los impactos de proyectos o programas (como un cambio de políticas, vidas que se salvaron o mejoraron, una mejora notable en las prácticas), se puede comprender con mayor profundidad el proceso de investigación y los factores del contexto que generaron estos resultados positivos o negativos en el mundo real.

Estas reflexiones tienen el potencial de mejorar la conducción de la investigación y de los investigadores hacia un cambio positivo de largo plazo. En el caso del IDRC, estas podrían ser algunas preguntas: ¿Las influencias clave utilizadas en el Marco de Referencia realmente inciden en la práctica (y cuáles son sus implicancias)? ¿Las subdimensiones de la calidad de investigación producen o mejoran la posibilidad de efectos de desarrollo? ¿La manera en que el IDRC enfoca la investigación para el desarrollo incentiva este tipo de investigaciones y produce impactos en el mundo real?



## **Tomar lo Aprendido para Mejorar el Uso de RQ+ en el Futuro**

Motivados por el aprendizaje y los resultados de la aplicación del RQ+ en sus Revisiones Externas, IDRC está considerando su uso potencial, no solo como mecanismo de evaluación, sino como marco de referencia para mejorar la gestión de las investigaciones. Como herramienta para este fin, el Marco de Referencia puede respaldar los proyectos y su ejecución, y ayudar a afianzar sus conceptos en los programas de IDRC y las investigaciones que financia.

## Referencias

- Bowen, S. & Graham, I.D. 2015. "Backwards design or looking sideways? Knowledge translation in the real world: Comment on 'A call for a backward design to knowledge translation'." *In J Health Policy Manag.* (2015). 4(8):545-547. DOI:10.15171/ijhpm.2015.71.
- Clark, W.C., Mitchell, R.B. & Cash, D.W. 2006. "Evaluating the Influence of Global Environmental Assessments." In Mitchell, R.B., Clark, W.C., Cash, D.W & Dickson, N.M. ed., *Global Environmental Assessments: Information and Influence*. Cambridge, MA: MIT Press.
- UK Department for International Development. 2014. *What is the Evidence on the Impact of Research for Development?* London, U.K.
- Eyre-Walker, A. & Stoletzki. 2013. "The Assessment of Science: The Relative Merits of Post-Publication Review, the Impact Factor, and the Number of Citations". *PLOS Biology*, October 8, 2013, DOI: 10.1371/journal.pbio.1001675.
- Feller, I. & Stern, P.C., ed. 2006. *A Strategy for Assessing Science: Behavioral and Social Research on Aging*. National Academy of Sciences, Washington, DC.
- Greenhalgh T. & Wieringa, S. 2011. "Is it time to drop the 'knowledge translation' metaphor? A critical literature review." *J.r. Soc. Med.* 104:501-9. DOI: 10.1258/jrsm.2011.110285.
- Guthrie, S., Wamae, W., Diepeveen, S., Wooding, S. & Grant, J. 2013. *Measuring Research: A Guide to Research Evaluation Frameworks and tools*. Prepared for the Association of American Medical Colleges. RAND Corporation, Santa Monica, CA, USA.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. & Rafols I. 2015. "The Leiden Manifesto for Research Metrics." *Nature*, April 23, 2015: 520: 429-431.
- International Development Research Centre. 2014. *Towards Research Excellence for Development: The Research Quality Plus Assessment Instrument*. Ottawa, Canada. Available at [idrc.ca/EN/Documents/Research-Quality-Plus-Assessment-Instrument.pdf](http://idrc.ca/EN/Documents/Research-Quality-Plus-Assessment-Instrument.pdf).
- Lavis, J., Lomas, J., Hamid, M. & Sewankambo, N.K. 2006. "Assessing country-level efforts to link research to action." *Bulletin of the World Health Organization*: 84:620-628.
- McNie, E.C., Parris, A. & Sarewitz, D. 2015. "A Typology for Assessing the Role of Users in Scientific Research: Discussion Paper." Paper prepared by the Consortium for Science, Policy and Outcomes at Arizona State University. Available at [cspo.org/wp-content/uploads/2015/02/Typology-Workshop-Paper.pdf](http://cspo.org/wp-content/uploads/2015/02/Typology-Workshop-Paper.pdf).
- Nowotny, H., Scott, P. & Gibbons, M. 2003. "Introduction. 'Mode 2' Revisited: The New Production of Knowledge." *Minerva* (2003) 41: 179-194.
- Nutley, S., Walter, I. & Davies, H.T.O. 2007. *Using Evidence*. The Policy Press, Bristol, UK.
- Ofir, Z. & Schwandt, T.S. 2012. *Understanding Research Excellence at IDRC: Final Report* (December 2012). Available at [idrc.ca/EN/Documents/RE-study-Understanding-RE-at-IDRC-full-report.pdf](http://idrc.ca/EN/Documents/RE-study-Understanding-RE-at-IDRC-full-report.pdf).
- Panel on Return on Investment in Health Research. 2009. *Making an Impact: A Preferred Framework and Indicators to Measure Returns on Investment in Health Research*. Canadian Academy of Health Sciences, Ottawa, Canada.
- Spaapen, J., Dijkstra, H. & Wamelink, F. 2007. *Evaluating Research in Context* 2nd, ed. Netherlands: Consultative Committee of Sector Councils for Research and Development, p. 12. Available at [qs.univie.ac.at/fileadmin/user\\_upload/qualitaetsicherung/PDF/Weitere\\_Aktivitaeten/Eric.pdf](http://qs.univie.ac.at/fileadmin/user_upload/qualitaetsicherung/PDF/Weitere_Aktivitaeten/Eric.pdf).
- Sarewitz, D. 2011. "Institutional Ecology and the Social Outcomes of Scientific Research". In Fealing, K.H., Lane, J., Marburger, J.H. & Shipp, S., ed., *The Science of Science Policy*, 337-348. Stanford University Press, Stanford, CA, USA.
- Stokes, D. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, DC: Brookings Institution Press.
- Wilsdon, J. et al., 2015. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Higher Education Funding Council of England, Bristol, UK.
- University of Gothenburg. 2013. "The Complex Role of Citations as a Measure of Scientific Quality," *Science Daily*, October 14, 2013.
- World Bank and Elsevier. 2014. *A Decade of Development in Sub-Saharan African Science, Technology, Engineering & Mathematics Research*.

