

**Análisis crítico de Estándares de Evaluación para América Latina y el Caribe, propuesta de la Red de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe, 2016<sup>1</sup>.**

**Critical analysis of the Evaluation Standards for Latin America and the Caribbean published by ReLAC, 2016.**

Tristán Ignacio Gramsch Calvo<sup>2</sup>  
tristangramsch@gmail.com

**Resumen**

En el presente artículo, se expondrán los resultados de una investigación sobre los Estándares de Evaluación para América Latina y el Caribe recientemente publicados gracias a una iniciativa de la Red de Seguimiento, Evaluación y Sistematización de América Latina y el Caribe (ReLAC). Mediante una entrevista en profundidad, los estándares fueron sometidos al juicio de un experto en el área de evaluación de políticas públicas. El artículo tiene como objetivos conocer los estándares de evaluación propuestos por la ReLAC, conocer cómo han sido integrados en las prácticas evaluativas en Chile, y cómo se vinculan con aspectos teóricos y metodológicos de las prácticas de evaluación en el país. Se concluye que, a pesar de la importancia de la inclusión de estándares de evaluación a las prácticas evaluativas chilenas, la generalidad que alcanzan los hacen inaplicables para todos los casos. Además, existen falencias en las prácticas evaluativas actuales que no permiten que parte de los estándares sean aplicados. En la primera sección se caracterizará brevemente al informante clave y se realizará un análisis de cada sección de los estándares de evaluación según su juicio. Subsiguientemente se indicarán aciertos, fortalezas y limitaciones de los estándares de evaluación, para finalmente realizar conclusiones.

Palabras clave: Estándares de evaluación, prácticas evaluativas, experto evaluación.

**Abstract**

In the present paper the results of an investigation about the Evaluation Standards for Latin America and the Caribbean recently published thanks to an initiative of the Monitoring, Evaluation and Systematization Network of Latin America and the Caribbean (ReLAC) will be shown. Through in-depth interview, the standards were subject to judgment by an expert in the area of public politics evaluation. The objective of the paper is to recognize the standards of evaluation proposed by ReLAC, recognize how these have been integrated in the evaluating practices in Chile, and how these are linked with theoretical and methodological aspects of the evaluating practices in Chile. It is concluded that, despite the importance that presents the inclusion of the standards in the evaluation practices in Chile, the broad aspect that covers make them inapplicable for all cases. Besides, there are flaws in evaluating practices that restrict the application of some of the standards. In the first section the key informant will be briefly described and an analysis of each section of the standards will be presented according to his

---

<sup>1</sup> La presente investigación se ha realizado en el marco de un estudio de mayor alcance, desarrollado en la cátedra de Evaluación de Proyectos, a cargo de la profesora Andrea Peroni, Facultad de Ciencias Sociales, Universidad de Chile

<sup>2</sup> Estudiante de Sociología, Universidad de Chile.

judgment. Then, successes, strengths and limitations of the standards will be shown that will lead to final conclusions.

Key words: Evaluation standards, evaluation practices, evaluation expert.

## **Introducción:**

Durante los últimos veinte años diferentes entidades alrededor del mundo han generado estándares de evaluación para guiar las prácticas profesionales de evaluación según principios más o menos generales (American Evaluation Association 2013) (Australian Evaluation Society 2013) (Canadian Evaluation Society 2011) (African Evaluation Association 2007) (Bilella, et al. 2016). Esto derivó en ciertos productos (American Evaluation Association 2004) (Australian Evaluation Society 2013) (Patel 2013) (The Canadian Evaluation Society 2010) que comparten rasgos comunes, a saber, la importancia otorgada a las relaciones interpersonales con los diversos actores, la rigurosidad técnica y analítica que debe alcanzar, la integridad y sistematización de la investigación, en otras. Sin embargo, América Latina y el Caribe no contaba con una experiencia similar que considerara la especificidad del territorio y sus prácticas evaluativas. Frente a tal déficit, ReLAC (Red de Seguimiento, Evaluación y Sistematización de América Latina y el Caribe) en conjunto con FOCEVAL (Proyecto Fomento a las Capacidades de Evaluación) y DEval (Agencia de Cooperación Alemana), publicaron Estándares de Evaluación para América Latina y el Caribe (Bilella, et al. 2016).

En el presente artículo se mostrarán los resultados de una investigación sobre tales estándares de evaluación, sometiéndolos al juicio crítico de un experto en el área. De esta manera, el artículo tiene como objetivos conocer en profundidad los estándares de evaluación propuestos por la ReLAC y cómo han sido integrados en las prácticas evaluativas en Chile. En el anexo se puede encontrar el instrumento de generación de información.

Se entrevistó a Víctor Salas el día siete de diciembre de 2016. Víctor Salas es académico, actualmente profesor titular de la USACH, especialista en áreas de evaluación a programas. Es ingeniero comercial graduado de la Universidad de Chile, tiene un magíster en Economía en la Universidad Católica de Lovaina y un doctorado en economía en la misma universidad. Ha participado en programas de evaluación del Ministerio de Hacienda desde hace más de diez años, realizando evaluaciones de programas gubernamentales

## **Estándares de evaluación:**

Con el incremento de las evaluaciones alrededor del mundo durante las últimas décadas (American Evaluation Association 2013) se han diversificado las estrategias metodológicas, las aproximaciones a los fenómenos evaluados, los alcances y destinatarios de las evaluaciones (Bilella, et al. 2016). Frente a este escenario, ReLAC, observando la experiencia en otros países, consideró necesario generar estándares que limitaran la forma en que se llevan a cabo las evaluaciones. Tales estándares son importantes para que las prácticas evaluativas tengan conclusiones más robustas y comparables, pero también para que las instituciones obtengan mejores resultados de sus políticas.

Los estándares de evaluación se dividen en cinco dimensiones de entre tres y ocho subdimensiones. Los hallazgos se presentarán siguiendo el orden de las dimensiones y sus subdimensiones, tal como se presenta a continuación. El análisis de la entrevista permite establecer si las dimensiones son cumplidas y aplicadas en las prácticas evaluativas o no, en base a la opinión y experiencia del experto.

Tabla 1<sup>3</sup>

Dimensiones y subdimensiones de los Estándares de Evaluación ReLAC

Dimensiones	Subdimensiones
<b>Evaluación rigurosa</b>	<i>Contextualizar la evaluación</i>
	<i>Descripción detallada del objeto de evaluación</i>
	<i>Preguntas evaluativas relevantes</i>
	<i>Metodología válida y confiable</i>
	<i>Niveles de participación adecuados de actores involucrados</i>
	<i>Conclusiones relevantes</i>
	<i>Recomendaciones útiles y factibles</i>
	<i>Reportes y comunicación pública efectiva</i>
<b>Adecuada evaluabilidad</b>	<i>Sobre gestión efectiva de la evaluación</i>
	<i>Procedimientos prácticos</i>
	<i>Viabilidad contextual, social y política</i>
	<i>Realista</i>
<b>Evaluación conducida de acuerdo con principios éticos y jurídicos</b>	<i>Respeto del derecho de las personas</i>
	<i>Autonomía</i>
	<i>Transparencia</i>
	<i>Legalidad</i>
<b>Adecuada comprensión cultural</b>	<i>Sobre igualdad y equidad</i>
	<i>Sobre derechos culturales</i>
	<i>Sobre reciprocidad</i>
<b>Relevancia y utilidad</b>	<i>Participación efectiva</i>
	<i>Propósitos acordados mutuamente</i>
	<i>Valores explícitos</i>
	<i>Información relevante, pertinente y oportuna</i>
	<i>Resultados útiles</i>
	<i>Comunicación y reportes puntuales y apropiados</i>
	<i>Interés por las consecuencias e incidencia</i>
<i>Valores expuestos</i>	

<sup>3</sup> Elaboración propia en base a Bilella (et al. 2016)

## **Hallazgos:**

### **Dimensión 1: Evaluación rigurosa**

El estándar de evaluación que esta dimensión propone (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de tener una serie de herramientas y cumplir una serie de procesos que la lleven a ser rigurosa, y por tanto, de mayor calidad y utilidad. De esa manera, para cumplir con esta dimensión se deben cumplir con los requisitos de sus subdimensiones. El análisis de la entrevista arroja que efectivamente esta dimensión es de gran importancia para la práctica evaluativa actual, y que gracias a esta pueden ubicarse las evaluaciones de calidad. Esto es, cumplir con las subdimensiones que se proponen lleva a una evaluación a ser considerada rigurosa. Sin embargo, esta dimensión tiene dos falencias empíricas, a saber, 1) que tiene una mirada que relega el economismo y 2) que no es lo suficientemente específica como para abordar todo tipo de evaluación.

1) Por una parte, la dimensión no tiende a considerar de mucha importancia la mirada economista de la evaluación, que se enfoca en el aporte que realiza la evaluación a la producción eficaz de bienes que produzcan innovaciones tecnológicas (cambios concretos en la realidad de la población). Por eso, la evaluación rigurosa, a pesar de que la ReLAC no proponga subdimensiones específicas respecto a esto (Bilella, et al. 2016), estaría definida por la ayuda que hacer para mejorar el proceso sistemático de producción de bienes. Esto requeriría de la capacidad del evaluador de indicar las falencias del programa evaluado.

2) El análisis anterior indica que existen una descoordinación entre lo que en la práctica evaluativa realiza y la propuesta de la ReLAC (Bilella, et al. 2016), especialmente, porque el contexto de aplicación de la evaluación hace que la dimensión de Evaluación Rigurosa tenga rendimientos limitados. En el caso de la experiencia de Víctor Salas la evaluación rigurosa está definida de manera más simple porque ha evaluado programas que deben lograr la generación de bienes. Por eso, el tipo de evaluación que se realiza, el tipo de programas evaluados, y la manera en que se aborda el problema hacen que las subdimensiones propuestas por la ReLAC (Bilella, et al. 2016) deban o no ser aplicadas.

#### **1.1 Contextualizar la evaluación**

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de caracterizar y comprender detalladamente el contexto local o regional en que la evaluación se desenvuelve. El análisis de la entrevista arroja que este estándar sí se cumple y se considera sumamente relevante. Sin embargo, el entendimiento de esta no concuerda completamente con el quehacer evaluativo actual, y requiere un análisis más detallado. Primero, la justificación requiere de la contextualización del programa, pero esta contextualización no pasa necesariamente por la descripción del contexto local (social, político, cultural y económico) de la población que está siendo evaluada. En el caso de la experiencia de Víctor Salas, la contextualización es de las innovaciones que han logrado realizar políticas y programas anteriores y el diagnóstico de evaluaciones anteriores de estas, buscando indicar si el programa debe existir o no. Esto es, la contextualización depende del programa que se está evaluando. Por eso, no se han realizado estándares de contextualización de evaluaciones, porque requerirían recursos y estudios específicos de caso.

## 1.2 Descripción detallada del objeto de evaluación

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como una descripción del objeto de evaluación lo suficientemente completa como para proponer un plan de cambio. El análisis de la entrevista indica que esta subdimensión puede o no incluirse en la evaluación, debido a que depende de la información que el programa evaluado tiene disponible. Esto se analizará a continuación. La descripción detallada del objeto de evaluación depende del grado de detalle de la información entregada en el programa y de la disponibilidad que de la contraparte de datos cruciales. Por ello, no es un estándar que la evaluación pueda cumplir de manera independiente. Sin embargo, en la experiencia de Víctor Salas, debido al trabajo con instituciones públicas funcionales, hay buen orden de la información y es fácil de ubicar. En específico, un problema de gestión importante que ha enfrentado el quehacer de la práctica evaluativa en Chile, dice de la incapacidad de los programas de generar indicadores para la medición de impactos. Claramente, esto afecta la producción de bienes para generar innovaciones. En general, estos se deben producir en la evaluación misma, imposibilitando una descripción detallada y obligando al evaluador a hacer algo más parecido a una descripción en conjunto.

## 1.3 Preguntas evaluativas relevantes

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la capacidad de la evaluación de hacer preguntas detalladas y pertinentes para los actores de la evaluación. El análisis de la entrevista indica que esta subdimensión sí se cumple y es relevante. La propuesta de matriz que aporta el programa es discutida por los evaluadores, que proponen nuevas aproximaciones para lograr de manera más eficiente y efectiva los productos (o para definir claramente productos). De esa manera, todos los componentes del programa se clarifican, logrando definir poblaciones poco difusas y metas pertinentes. Si las preguntas son relevantes, lograrán aportar al programa en estas áreas. Sin embargo, la subdimensión de preguntas relevantes no enfatiza la mirada economista, a saber, que las preguntas deben enfocarse en los objetivos del programa, para lograr evaluaciones que no son puramente formales.

## 1.4 Metodología válida y confiable

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse en que el diseño de la evaluación y sus procedimientos de recolección de datos deben ser rigurosos, y orientados según los propósitos evaluativos. El análisis de la entrevista indica que sí se cumple esta dimensión y es relevante. En específico, una metodología confiable está constantemente observando sus propias prácticas, lo que le permite, por ejemplo, no forzar una metodología inapropiada en la evaluación. En esta misma línea, Víctor Salas indica que la evaluación siempre debería tener un grado de flexibilidad metodológica suficiente como para enfrentar la especificidad del programa, que puede tener datos difusos, estar desordenada, confundir conceptos, etc. Sin embargo, esta subdimensión no establece cuáles deberían ser los propósitos de la evaluación y, por ello, tampoco indica que la metodología debería apuntar a cumplir ese propósito. Para el entrevistado, estos deberían girar siempre en torno a bienes. Así, una metodología válida de evaluación es aquella que se adecúa a las necesidades del programa para ayudar a generar o redefinir los bienes que propuso producir.

## 1.5 Niveles de participación adecuados de actores involucrados

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la importancia que la evaluación de a todos los actores influidos por la evaluación, mostrando las diferentes perspectivas involucradas y la importancia de los diferentes actores. El análisis de la entrevista indica que esta dimensión no se cumple en la práctica evaluativa del entrevistado porque no está en contacto con los actores de la población afectada por los programas que evalúa. Es difícil integrar a los actores beneficiarios del programa en la evaluación porque hay poco contacto directo con estos. La participación, en la práctica evaluativa actual, está limitada por el tipo de evaluación que se realice. Y las veces en que interactúa más estrechamente con los actores beneficiarios del programa, ha sido para generar información necesaria para la evaluación. Sin embargo, no es sólo un problema de la evaluación que los actores no sean integrados (y de esto no da cuenta el estándar) ya que los programas tienden a ser poco participativos e incluir poco en su diseño a los usuarios de los programas. Sumado a esto, la poca claridad de muchos programas hace que la evaluación no pueda incluir a los actores porque no han sido claramente definidos, y es la evaluación misma la que debe encargarse de ello. En ocasiones, la evaluación llega delimitar a los actores involucrados y, evidentemente, no puede reconocerlos de antemano.

#### 1.6 Conclusiones relevantes

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de que la evaluación tenga conclusiones empíricamente fundadas y en concordancia con sus propósitos de evaluación. El análisis de la entrevista arroja que esta dimensión sí se cumple pero no concuerda con precisión con la postura de Víctor Salas. Para el entrevistado las conclusiones relevantes son aquellas que pueden traducirse en innovaciones tecnológicas que permitan cambiar la realidad de una población. Esto conlleva que no deben aportar únicamente a la literatura científica, aunque puedan hacerlo. Por ello, una conclusión relevante requiere de procedimientos que permitan que sean claras y bien fundadas, pero logra ser relevante sólo si genera la innovación. En la práctica profesional en Chile, muchas evaluaciones pueden no acabar en innovaciones, lo cual es una falencia que debe ser superada.

#### 1.7 Recomendaciones útiles y factibles

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de ofrecer recomendaciones útiles y posibles de implementar gracias a la participación de la población intervenida. El análisis de la entrevista arroja que esta dimensión sí se cumple pero tiene falencias de aplicación a ciertos contextos. Por una parte, sí se cumple porque las recomendaciones están hechas gracias a procedimientos evaluativos rigurosos, basados en los resultados que ofrecen mejoras factibles. Sin embargo, la participación de la población beneficiaria de los programas evaluados no está siempre asegurada, porque las evaluaciones se realizan sólo de los programas. Las recomendaciones, en general, se han enfocado en revisar las decisiones del programa y aportar con argumentos para su continuación o cancelación según los resultados que ha obtenido. Pero, a su vez, las recomendaciones no siempre logran ser útiles porque la contraparte no le da la importancia que merecen. Por otro lado, las recomendaciones están hechas dependiendo de los casos, no en base a criterios predefinidos. Esta flexibilidad es necesaria cuando se tiene en cuenta de que criterios estandarizados pueden dejar de medir la eficacia de un programa. Si tuviese que aplicarse un criterio, debería ser que se están financiado programas que generan innovaciones tecnológicas, este podría ser un protocolo de eficiencia necesario, que aseguraría el relego de programas no

eficaces. Por todo lo anterior, puede observarse que las recomendaciones útiles y factibles no sólo dependen de la capacidad de la evaluación, sino de la contraparte también.

### 1.8 Reportes y comunicación pública efectiva

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de comunicar efectivamente los resultados de evaluación a los actores involucrados y público general, realizando informes que puedan ser públicamente leídos (si no violan la confidencialidad de las partes). El análisis de la entrevista arroja que la subdimensión sí se cumple y resulta de gran importancia su comunicación efectiva, a pesar de que en ocasiones la contraparte no le da la importancia que merece. Los programas que Víctor Salas ha evaluado tienden a saber a quién debe realizarse la comunicación de los resultados evaluativos: el parlamento. En su experiencia, los resultados han sido expuestos a expertos en el área, y los reportes han considerado todos los aspectos evaluados, incluso aquellos nacidos de la conversación. Sin embargo, a pesar de su contacto con expertos, han habido desencuentros y malos entendidos que requirieron de una comunicación constante entre las partes para resolverse. Por ello, en general, la comunicación efectiva con la contraparte no ha sido un problema. Pero a pesar de las condiciones específicas de comunicación pública efectiva, aún existen problemas de difusión que el evaluador no controla, a saber, que la contraparte no siempre lee ni revisa los documentos generados por la evaluación.

### **Dimensión 2: Adecuada Evaluabilidad**

El estándar de evaluación que esta dimensión propone (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de cumplir con requisitos de eficiencia y eficacia para lograr una evaluabilidad prudente y consciente de los costos que implica. Para esto, debe existir una fase de diseño que asegure que se establezcan criterios alcanzables. El análisis de la entrevista arroja que efectivamente esta dimensión es importante en la práctica evaluativa actual. Pero, 1) presenta un grado de generalidad que la hace sólo parcialmente adecuada para ser estándar de evaluación de todas las evaluaciones de proyectos. Y 2) no depende únicamente de las prácticas evaluativas, sino aún más de los programas evaluados.

1) La generalidad de las consideraciones y definiciones de toda la dimensión, hacen que una adecuada evaluabilidad sea propia de evaluaciones que trabajan en condiciones específicas, por ejemplo, de rigurosidad de la contraparte en el contacto. Por otro lado, ciertos criterios establecidos por la dimensión son amplios y no se traducen en condiciones concretas que debe tener una evaluación, lo que implica que en cada evaluación las formas de medirlo variarán.

2) El panorama de la práctica evaluativa actual muestra que existen, de antemano, ciertos programas que serán evaluados y otros que no porque la magnitud de su intervención lo amerita. Por ello, la condición de evaluabilidad de los programas viene dada y el diseño de la evaluación depende de estos. No todos los programas están pensados para ser evaluados y, por ello, tienden a no cumplir con los requisitos más estrictos de aquellos que sí serán evaluados. Por ello, una adecuada evaluabilidad también depende del programa. En el caso de que el programa no esté bien estructurado y tenga falencias de diseño, la evaluabilidad será más forzada, al menos con metodologías complejas, como la de marco lógico.

### 2.1 Gestión efectiva de la evaluación

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de contar con estrategias eficientes y efectivas, y utilizar bien los recursos. El análisis de la entrevista arroja que esta subdimensión sí se cumple y es importante. Víctor Salas coloca especial énfasis en que la gestión efectiva de evaluación no sólo debe contar con procedimientos rigurosos, sino también, debe generarse en cooperación con todo el equipo evaluador. Se debe lograr “amasar” la gestión de la evaluación. Debido a que el equipo de evaluación se arma en el momento en que debe realizarse la evaluación, el grupo de expertos que lo compone no se conoce entre sí. En reuniones se establecen los puntos importantes de cada evaluación y cómo se realizará su gestión. Se revisan los elementos que son susceptibles de evaluación, siendo los bienes aquellos más importantes. Por eso, la gestión efectiva de la evaluación depende en gran parte de la gestión de la contraparte y de si tienen la información necesaria para hacerla.

## 2.2 Procedimientos prácticos

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de ajustar sus procedimientos a las prácticas de evaluación. El análisis de la entrevista arroja que esta subdimensión sí se cumple pero que debe considerar la especificidad que enfrenta cada evaluación. Esto porque los procedimientos de evaluación se adecúan al objeto de estudio, que a su vez depende del programa evaluado. Si la contraparte no está dispuesta a entregar cierta información, a facilitar datos, a proporcionar contactos, los procedimientos evaluativos se ven entorpecidos. En el caso de que sean reacios a entregar información, el evaluador debe buscar trabajar con la contraparte y lograr obtener lo necesario. En términos prácticos, el evaluador debe tener la flexibilidad para acercarse a la contraparte e indicarle lo que esta no ve, por ejemplo, de que sí hay datos con los cuáles trabajar pero que requieren ser filtrados. De hecho, Víctor Salas ha tenido experiencia con bases de datos requieren del trabajo de los evaluadores para ser utilizadas. Así, de manera práctica, la evaluación es con la contraparte, y no sólo de la contraparte; sólo así el evaluador puede adentrarse en las prácticas de aplicación del programa.

## 2.3 Viabilidad social, contextual y política

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de las evaluaciones de reconocer e incluir intereses y necesidades, culturales y políticas, de los diversos actores. El análisis de la entrevista arroja que esta subdimensión no se cumple porque, en general, las evaluaciones que realiza Víctor Salas no requieren entrar en contacto con la población beneficiaria del programa. Así, la evaluación de la viabilidad contextual, social y política depende del programa que se esté estudiando, y mayoritariamente son los programas que tratan con poblaciones vulnerables los que lo requieren. Además, es especialmente difícil contextualizar la evaluación si los beneficiarios no están presentes. Evaluaciones de esta naturaleza, que no están en contacto con la población beneficiaria, no pueden incorporar a todos los actores en el proceso. El quehacer de la evaluación chilena, sin embargo, se realiza en pos del beneficio de estos actores a pesar de que no estén incluidos en la discusión de la evaluación. Por ello, la evaluación no puede siempre incluir todos los intereses culturales en juego.

## 2.4 Realista

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de tener una clara coherencia entre los recursos disponibles, el acceso a estos, y los propósitos de evaluación. El análisis muestra que esta subdimensión no se cumple porque existen varios factores que impiden el adecuado quehacer de la evaluación y que no dependen únicamente de esta. De hecho, esta subdimensión es particularmente difícil de abordar porque comúnmente se trabaja con programas que siguen operando. En la marcha, es difícil convencer a la contraparte de cambiar o volver pensar la ejecución del programa, en especial, porque están saturadas las funciones y no tienen personal al que pueda atribuírsele más carga laboral. Los programas no pueden dedicarle demasiado tiempo a la evaluación y, por ello, no siempre se tiene el efecto esperado con la novedad introducida por la evaluación, a saber, los resultados respecto a los bienes que deben producirse. Aparte de esto, para Víctor Salas el trabajo de evaluador no entrega exclusiva atención a la evaluación. Comúnmente se trabaja en otras cosas al mismo tiempo. Es un proceso de poco tiempo, en el que el punto de partida es poco claro si no se ocupan las metodologías adecuadas (como de marco lógico). Una evaluación realista, por lo tanto, debe considerar que tendrá que enfrentarse en muchas ocasiones a poco tiempo y muchas tareas. Como vemos, no sólo depende de la planificación de la evaluación, sino también, del quehacer en la evaluación misma.

### **Dimensión 3: Evaluación conducida de acuerdo con principios éticos y jurídicos**

La dimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como el aseguramiento de que en el proceso evaluativo se reconozcan los derechos de los actores involucrados, garantizando su respeto y sus responsabilidades. El análisis muestra que esta dimensión se cumple parcialmente, porque hay subdimensiones que no se cumplen. En específico, el tipo de evaluación que parte de los profesionales en Chile realiza no requiere ser conducida respetando el derecho de los actores porque tiene una posibilidad muy baja de vulnerarlos. Y, además, los principios jurídicos están presentes de antemano en la evaluación porque las instituciones públicas que solicitan las evaluaciones los cubren. Si es que llegan a aparecer situaciones que parecen ser legalmente inapropiadas se le informa oportunamente al organismo que se encarga de la evaluación. Sin embargo, la evaluación para Víctor Salas sí debe tener un principio ético fundamental, a saber, la autonomía de los evaluadores respecto del organismo que encarga la evaluación. No se debe ni hacer ni decir lo que el organismo quiera. Pero es necesario aclarar que la autonomía no conlleva ignorar los requisitos de evaluación del organismo, como la justificación de resultados. De esa manera, esta dimensión es parcialmente cumplida.

#### **3.1 Respeto del derecho de las personas**

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de proteger el derecho de los actores involucrados, siendo sensible a sus creencias y usando el consentimiento informado. Puede considerarse que esta subdimensión no aplica en el esquema de si se cumple o no, porque no existen las condiciones para su realización en ciertas prácticas evaluativas. En específico, el resguardo de los derechos de los actores involucrados está asegurado durante la fase de acercamiento obligatorio, y no están resguardados por los evaluadores, sino, por la institución que solicita la evaluación. Pero en instancias posteriores, no se conocen los derechos que deben respetarse de los involucrados. En la fase de acercamiento obligatorio, se conoce que los datos proporcionados para la evaluación no deben usarse fuera de esta para no vulnerar su confidencialidad, pero no hay procedimientos

estándar para ello. La cercanía con quienes manejan las bases de datos podría estar vulnerando derechos pero no se sabe. Las limitaciones generadas por la legalidad han sido comúnmente superficiales, como que no se comparten bases de datos, situaciones superadas gestionando soluciones con la organización que solicita la evaluación. Como se puede ver, hay evaluaciones que no pueden vulnerar los derechos de los actores, porque no están involucrados con ellos.

### 3.2 Autonomía

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de que la evaluación y sus procedimientos sean independientes de otras funciones institucionales, cosa de que su imparcialidad y transparencia no se vean afectadas. El análisis de la entrevista muestra que esta subdimensión sí se cumple y es de extrema importancia para la evaluación. La autonomía es cautelada por el organismo mismo que solicita la evaluación, gracias a un concurso público que permite que se elijan proyectos de evaluación y no a personalidades específicas. Pero a medida que los ejecutores de programas van conociendo a los evaluadores pueden saber más o menos cuál es su perfil, lo que podría derivar en que se elija a un evaluador sobre otro, por ejemplo, por la conocida dureza de sus críticas. El evaluador necesita ser autónomo sin dejar de responder a los propósitos de la evaluación de la contraparte ni a los propios. Esto se logra gracias a la libertad de elegir cómo operar en los diversos contextos de evaluación, de hecho, las formas de operar no necesitan ser comunicadas al organismo que solicita la evaluación, resguardando la autonomía. Víctor Salas indicó: "Nosotros no tenemos que hacer y decir lo que la DIPRES quiera, eso yo lo defiendo a muerte, tenemos autonomía en la evaluación."

### 3.3 Transparencia

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de generar documentos públicos disponibles para las partes y actores interesados, a menos que se deba resguardar su confidencialidad. El análisis de la entrevista muestra que esta dimensión sí se cumple pero depende de la rigurosidad de la evaluación. Si la evaluación es adecuada y contiene la cantidad de componentes necesarios para ser rigurosa, logrando tener conclusiones fundamentadas empíricamente, entonces es positivo publicarla. Para Víctor Salas, Chile ha avanzado bastante en esta materia durante los últimos años. Sin embargo, con respaldo en su experiencia, considera que los trabajos de evaluación deberían siempre porque son financiados por el Estado, por lo que la sociedad debería conocer si los recursos fueron bien utilizados por los evaluadores.

### 3.4 Legalidad

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de estar respaldada por instrumentos legales, firmados entre las partes, que expliciten acuerdos, delimitando detalladamente los componentes, costos y funciones, de la evaluación. El análisis muestra que esta dimensión no se cumple, debido a su falta de necesidad en el proceso evaluativo. En específico, la legalidad de la evaluación no es preocupación de los evaluadores porque el organismo que solicitó la evaluación cubre estos aspectos. Son los programas los que requieren de una práctica auto observadora de legalidad, facilitada por el organismo que los solicita. De esa manera, los evaluadores llegan a evaluar un programa que es legal de antemano. Además, los instrumentos legales que coordinan a las partes,

están de antemano asegurados por el organismo que solicita la evaluación. Por ello, esta subdimensión no se cumple, ya que la legalidad está definida antes del proceso evaluativo.

#### **Dimensión 4: Adecuada comprensión cultural**

La dimensión de Adecuada comprensión cultural propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de respetar y ser sensible a los contextos históricos y culturales de la población que evalúa, asegurando una comunicación efectiva con ellos. El análisis de la entrevista arroja que esta dimensión no se cumple porque ninguna de sus subdimensiones se cumple: la especificidad de ciertas evaluaciones hace que el contacto con la población beneficiaria de programas sea escasa o nula, no pudiendo adecuarse a su contexto cultural. Sin embargo, es rescatable una idea importante de esta dimensión, a saber, que a pesar de que no hay contacto con la población beneficiaria de los programas, la evaluación debe comprender adecuadamente a los ejecutores del programa. Así, las evaluaciones deben ser sensibles a los diferentes contextos culturales que enfrenta, en especial, porque los programas pueden ser llevados a cabo por una multiplicidad de actores distintos y de diversas procedencias. Además, los ejecutores del programa pueden variar a través del tiempo y conocer la evolución del uso de términos o de objetivos en el programa es importante para comprenderlo actualmente. Las definiciones de los conceptos principales, como de objetivos y propósitos, pueden ser nebulosas incluso dentro del mismo programa. Tanto así que, por ejemplo, Víctor Salas ha tenido que recurrir a ejecutores pasados para comprender qué quisieron decir con algunas partes del programa.

Por ello, si los paneles se componen por expertos que dominan la metodología y el sector geográfico con el que trabajan, siempre habrá alguien que podrá realizar el intercambio en el lenguaje correcto con la población objetivo, logrando una adecuación cultural necesaria para la buena ejecución del programa. Sin embargo, la dimensión de "Adecuada comprensión cultural" refiere a que las evaluaciones deben ser sensibles a los contextos culturales de las poblaciones afectadas por la evaluación, y en el caso de algunas evaluaciones en Chile, esto tiende a no ocurrir porque no está en contacto con estas. Así, debe enfrentar la complejidad cultural del equipo ejecutor, que puede llegar a ser un gran desafío, pero no la de la población afectada.

##### **4.1 Igualdad y equidad**

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de incluir a todos los actores involucrados sin importar su condición o jerarquía, contribuyendo a la tolerancia para estos. El análisis de la entrevista arroja que esta subdimensión no se cumple porque parte de las evaluaciones no trabajan directamente con la población beneficiaria del programa. En específico, Víctor Salas no realiza casi nunca evaluaciones en que deba trabajar directamente con la población. Así, los programas sobre los cuales realiza evaluaciones no siempre requieren del contacto con los beneficiarios del programa. Sin embargo, en los casos en que sí se requirió, el experto supo trabajar respetuosa y equitativamente con la población objetivo. De esta manera, puede observarse que no se cumple esta subdimensión porque no se trabaja con los actores beneficiados y, por eso, no pueden incluirse equitativamente en la evaluación; lo que no quita que sea importante hacerlo cuando sí estén incluidos.

##### **4.2 Derechos culturales**

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de que las evaluaciones se realicen respetando las identidades culturales y dignidad de los actores, gracias a protocolos de atención adecuados. El análisis muestra que esta subdimensión tampoco se cumple porque, nuevamente, no existen las condiciones de contacto con la población beneficiaria de los programas. Y, además, la necesidad de un protocolo para tratar con las diferencias culturales de la población sólo debería existir si se tratan poblaciones grandes que lo requieran. En comunidades pequeñas, la flexibilidad para aproximarse puede generar mejores resultados, debido a un manejo más libre de la relación. En específico, esto se da porque comúnmente no se aplican los mismos instrumentos de medición a poblaciones de distintos tamaños. Víctor Salas considera que no se realizan demasiadas evaluaciones de impacto, que sí requerirían el uso de estas herramientas a gran escala y protocolos de adecuación cultural. Por lo tanto, establecer protocolos culturales no es necesario para la realización de cierto tipo de evaluaciones.

#### 4.3 Reciprocidad e identidades culturales

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de las evaluaciones de tener relaciones recíprocas con los actores involucrados, respetando valores y modos de pensar, logrando mayor cohesión y confianza con la contraparte. El análisis de la entrevista indica que esta subdimensión no se cumple porque algunas evaluaciones no tienen la posibilidad de entablar este tipo de relaciones con la población beneficiaria. En específico, Víctor Salas no ha tenido experiencias significativas en que haya debido lograr reciprocidad entre los actores involucrados ni entre identidades culturales diversas. En una ocasión, tuvo que hablar con altos funcionarios de varias universidades y jefes de proyectos de estas, entablando una relación respetuosa, pero no tuvo que realizar consideraciones específicas de reciprocidad entre culturas o grupos humanos. Esta subdimensión no aplica porque, sin importar a qué alude, ciertas evaluaciones no pueden establecer el vínculo que propone como recomendable.

#### **Dimensión 5: Relevancia y utilidad**

El estándar de evaluación que esta dimensión propone (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de ser relevante para los actores involucrados, aportando en las decisiones que toman y logrando satisfacer sus necesidades. El análisis de la dimensión arroja que sólo parte de sus dimensiones se cumplen, porque algunas no aplican para todos los tipos de evaluación realizadas. En específico, porque las evaluaciones útiles y relevantes para el programa no siempre deben cumplir con los estándares de las subdimensiones que propone la ReLAC. La relevancia y utilidad de una evaluación está definida por, principalmente, su capacidad de entregar resultados claros y precisos. Estos deben indicar si se están o no produciendo eficazmente bienes que introduzcan innovaciones tecnológicas en la población que atiende el programa.

Víctor Salas, de manera crítica, observa una carencia que los estándares de evaluación no están considerando, a saber, que la información entregada por los evaluadores se usa poco y, por lo tanto, la relevancia y la utilidad de esta dependen de la contraparte. El interés por los resultados de evaluación es poco y no hay procedimientos formales para que todos los ejecutores de programas conozcan estos resultados. En específico, considerando la experiencia del entrevistado, las empresas que han visto sus productos evaluados no observan los productos de la

evaluación. Simplemente no se sabe si leen los reportes. Además, tampoco hay mecanismos para saber si agentes que solicitan las evaluaciones conocen los resultados de estas.

### 5.1 Participación efectiva y consciente

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de que todos los actores involucrados participen activamente. El análisis de la entrevista arroja que esta dimensión no se cumple porque algunas evaluaciones no tienen contacto con todos los actores que involucra, y porque también depende de la contraparte. La participación de los actores, para Víctor Salas, se da cuando la contraparte está efectivamente interesada en ser parte del proceso evaluativo. Para eso, deben conocer lo más profundamente posible los estudios de evaluación realizados anteriormente, y deben conocer la cantidad de documentos que se han generado. Sin embargo, no siempre la contraparte está interesada en el proceso de evaluación ni tampoco conoce en profundidad los documentos que necesita conocer. Por ello, la participación de los actores no está asegurada porque son los mismos evaluados los que no están interesados en la evaluación. Por otro lado, la participación de otros actores, como de la población, no está asegurada porque no hay contacto con los investigadores en todo tipo de evaluaciones.

### 5.2 Propósitos acordados mutuamente

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de que la evaluación sea un proceso acordado con la contraparte, en todos deben contribuir para lograr una evaluación precisa y completa. El análisis de esta subdimensión arroja que el estándar sí aplica, debido a que ambas partes deben establecer una comunicación constante que los lleve a la aclaración efectiva de los propósitos de la evaluación. De hecho, los posibles intermediarios, como los que solicitan la evaluación, pueden aportar a que los partes lleguen a propósitos mejor acordados en conjunto, por ejemplo, mediante el establecimiento de información necesaria en una matriz que todos conocen. Esto lograría una buena evaluación en tanto entrega todos los insumos que requiere. Y en respecto a la organización interna de los evaluadores, la realización de reuniones de reuniones para acordar qué se va a evaluar en base a la información disponible, es de gran relevancia.

### 5.3 Valores explícitos

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de exponer los valores que guían los juicios del evaluador. El análisis de la entrevista arroja que esta subdimensión no se cumple, debido a la falta de instancias para tal explicitación. Además, los juicios del evaluador no parecen ser demasiado importantes porque los programas ya muestran de manera imparcial toda la información necesaria. Así, si se conoce bien el programa, la aplicación de metodologías suele ser estandarizada. Sólo en la sección de recomendaciones se emiten juicios de valor explícitamente, pero no se atribuyen al evaluador. A pesar de lo anterior, Víctor Salas da cuenta de que los expertos en evaluación tienen juicios de valor anteriores a la evaluación misma. Por ello, sus evaluaciones sí dan cuenta de que existen juicios de valor que afectan la investigación, pero siempre existirán, y su explicitación no parece necesaria. Sin embargo, una posible subdimensión sea necesaria aquí, a saber, una que establezca específicamente dónde explicitar los juicios de valor. Se ahondará en esto más adelante.

#### 5.4 Información relevante, pertinente y oportuna

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de ser útil a la necesidades de los actores y grupos involucrados en la evaluación. El análisis de esta subdimensión arroja que sí se cumple, a pesar de problemas de poca lectura de informes evaluativos. Esto porque la información generada en las evaluaciones ha sido altamente importante para la toma de decisiones institucionales de la contraparte. Víctor Salas, en sus experiencias positivas, ha tenido un diálogo fluido con la contraparte que indica que hay un gran interés por conocer el trabajo y los cambios que puede generar. De hecho, conoce de muchos programas que trabajan con la información de su evaluación. Lo anterior deriva en que los evaluados cambian sus prácticas debido a la evaluación y, también, cambian los programas, que se ven confrontados a necesidades que no observaban. Sin embargo, Víctor Salas también observa que se han hecho asesorías valiosas por expertos que no han podido ser publicadas y que finalmente no son traducidas a textos de evaluación. Esto es, información relevante sobre programas es finalmente desechada porque no se traduce en textos.

#### 5.5 Resultados útiles

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad del programa de construir resultados que animen y permitan a los actores involucrados observar de nuevo sus prácticas respecto al objeto evaluado, con una nueva perspectiva. El análisis de esta subdimensión arroja que sí se cumple porque, a pesar de la entrega de resultados que no pueden ser aplicados como bienes que produzcan innovaciones tecnológicas, la mayoría de estos lleva a la contraparte a un cambio. Víctor Salas, sabe que sus resultados han sido útiles porque sigue en contacto con muchos de los programas que evaluó y observa cómo han cambiado según su evaluación. No obstante, muchas de las evaluaciones que él conoce (y algunas de las cuales fue parte) no se traducen en cambios ni en el programa ni en el quehacer de la contraparte. Esto significa que, a pesar de que expertos generen información útil y relevante para el programa evaluado, esta no siempre genera cambios importantes en la contraparte. En rigor, esta subdimensión es cumplida por el entrevistado, pero debe considerarse que no es así en todos los casos.

#### 5.6 Comunicación y reportes puntuales y apropiados

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación de comunicarse efectivamente con las diferentes audiencias mediante reportes breves y apropiados. El análisis de la entrevista arroja que esta subdimensión sí se cumple y que no presenta dificultades en el contexto de las evaluaciones revisadas. La comunicación de los resultados de investigación evaluativa ha tendido a no presentar demasiadas dificultades por estar pensada para expertos que manejan los tecnicismos y conocen el rubro de trabajo. De esa manera, Víctor Salas ha tenido experiencias en que ha sido sumamente claro con sus reportes, que logran cambiar e incluso acabar con algunos programas. Esto influye claramente en la vida de los actores que deben lidiar con los posibles cambios, desde los ejecutores del programa hasta los beneficiarios de este.

#### 5.7 Interés por las consecuencias e incidencia

La subdimensión propuesta por la ReLAC (Bilella, et al. 2016) puede resumirse como la necesidad de la evaluación por interesarse en el bienestar de los actores participantes en la evaluación, promoviendo el buen uso de la información producida, y buscando evitar que actos tengan efectos negativos sobre los actores. El análisis de la subdimensión arroja que sí se cumple porque se busca la generación de bienes que finalmente cambien la realidad de la población que atiende el programa. Para ello, el panel debe tener la capacidad para generar un diagnóstico que ubique las falencias del programa, logrando que elimine las trabas a la producción o distribución de bienes. Así, la evaluación siempre debe derivar en la generación de innovaciones tecnológicas para el beneficio de la población atendida. El interés de los evaluadores es de finalmente generar un impacto en la población carenciada, mejorando su nivel de vida, o permitiéndoles obtener aquello de lo que carecen.

### **Aciertos y fortalezas de los estándares de evaluación**

El análisis de la entrevista permite indicar aciertos y desaciertos de los estándares de evaluación. Tal indicación, a su vez, permite agregar otras subdimensiones necesarias en los estándares.

Las fortalezas que presentan estos estándares de evaluación pueden resumirse en dos grandes puntos, a saber, *a)* su capacidad de abarcar una serie larga de evaluaciones distintas, *b)* la preocupación por la inclusión efectiva de todos los actores que participan en la evaluación, *c)* la posibilidad de ubicar empíricamente estos estándares en las prácticas evaluativas chilenas.

*a)* La generalidad de los estándares de evaluación permite que, sin importar la especificidad de la evaluación que se esté observando, existen herramientas para juzgar si es completa o no. Esto es un claro avance en el ejercicio de cohesión de la práctica evaluativa latinoamericana. Y, además, es una forma de discernir las evaluaciones rigurosas de aquellas que no lo son. Permite que los evaluadores conozcan qué es lo que otros evaluadores consideran necesario que tenga un proyecto de evaluación. De esta manera, lograría, si efectivamente tiene impacto en las prácticas evaluativas, que una serie larga de evaluaciones distintas se rijan o al menos consideren ciertos estándares de evaluación.

*b)* A pesar de que, en general, es intuitivo para los expertos que los estándares que establece la ReLAC deben cumplirse, no es intuitiva la importancia que debe dársele a los diversos actores que son parte de la evaluación. Bilella et al. (2016) lo considera necesario delante del contexto latinoamericano de gran diversidad cultural y étnica y, por eso, da especial importancia al resguardo de la diversidad y el respeto por el otro. Conocer los rasgos propios de las evaluaciones realizadas en un territorio, es fundamental para establecer diferencias útiles en la investigación. Que los estándares se enfoquen en la especificidad latinoamericana y que hagan de su inclusión en la evaluación es una fortaleza de los estándares.

*c)* A pesar de que no se leyeron completamente los estándares a Víctor Salas, encontró que en la mayoría de estos su práctica profesional y la de muchas evaluaciones en Chile está reflejada. Esto significa que los estándares aciertan en dar cuenta de gran parte de las prácticas típicas en países latinoamericanos.

Un juicio crítico de estos estándares, que intuitivamente pareciera que deben cumplirse siempre, permite indicar debilidades que deben considerarse. La más importante de estas, es que *1)* la generalidad de los estándares no permite dar cuenta de las especificidades de las evaluaciones en

contextos institucionales que no requieren del contacto con todos los actores. Además, 2) el enfoque economista está dejado de lado porque no se da énfasis a la importancia que tiene para la evaluación generar bienes que produzcan innovaciones tecnológicas que finalmente mejoren el bienestar de la población beneficiada por el programa. Por último, 3) los estándares no contienen consideraciones necesarias respecto a exigencias a la contraparte para que se realice adecuadamente el proceso de evaluación.

1) La especificidad de las evaluaciones realizadas por Víctor Salas y por otros evaluadores en Chile, hacen que no estén en contacto con todos los actores de la evaluación. Esto hace que una parte importante de las subdimensiones propuestas por la ReLAC no puedan ser incluidas en estas prácticas evaluativas. Por ejemplo, aquellas que hablan del respeto y la integración de todos los actores a la evaluación. Por ello, la generalidad de las dimensiones hace que su cumplimiento no signifique necesariamente algo positivo o deseable: en algunos casos simplemente no puede darse.

2) Los estándares de evaluación tienden a no prestar demasiada atención al enfoque economista, que enfatiza la necesidad de la evaluación de aportar en la generación y distribución de bienes del programa para cumplir los propósitos de este. Esto es una falencia cuando la práctica de evaluación ha mostrado que el enfoque economista es efectivo en traer lograr beneficios a la población y cumplimiento de propósitos del programa.

3) Los estándares no contienen consideraciones importantes respecto a exigencias a la contraparte para la realización de un adecuado proceso evaluativo. Por una parte, la información entregada por los evaluadores se usa poco y no hay estándar que regule esto. El interés por los resultados de evaluación es poco y no hay procedimientos formales para que todos los ejecutores de programas conozcan estos resultados. En específico, en la experiencia de Víctor Salas, las empresas que han visto sus productos evaluados no observan los productos de la evaluación. Simplemente no se sabe si leen los reportes. Además, no hay mecanismo para saber si hasta los agentes que piden las evaluaciones conocen los resultados de estas. Si es que no se consideran los resultados entregados a la contraparte, no importa si son útiles o relevantes, sigue siendo información generada que no tendrá efectos en las prácticas de ejecutores de programas.

Así, no hay estándares de revisión de información producida, lo que hace que no se sepa quién mira la evaluación realizada. Muchos estudios terminan "guardándose en un fichero" para ser posteriormente olvidados. De hecho, Víctor Salas sugiere que los ministros y los equipos que toman decisiones presupuestarias, como personas de altos rangos en los ministerios, deberían conocer el resultado de las evaluaciones.

Por otra parte, no hay estándares que regulen la comunicación constante de la contraparte con los evaluadores. Cuando los programas se encuentran en momentos resolutivos, como cierres de período, los evaluadores pueden dejar de ser atendidos o pueden verse atendidos de manera parcial. En ocasiones el tiempo es escaso y esto afecta el proceso evaluativo. Además, tampoco hay estándares de cooperación informativa de parte de la contraparte, lo que significa que no siempre se obtendrá la información necesaria para hacer una buena evaluación.

Si se llegara a argumentar que estas consideraciones no pasan por el dominio de la evaluación, se debe entender que muchos de los estándares propuestos por la ReLAC tampoco, porque

dependen parcial o completamente de la acción de la contraparte. Una evaluación rigurosa requiere también de la cooperación rigurosa de la contraparte.

Por todo lo anterior, la presente investigación considera necesario que se agregue una nueva dimensión a los estándares de evaluación propuestos por la ReLAC, a saber, Aseguramiento de disposición de la contraparte. Esta se definiría como la necesidad de los evaluadores de establecer, previo al proceso de implementación de la evaluación, que la contraparte entregará la información necesaria para la evaluación, será constante en las comunicaciones con esta, y dispondrá parte de su tiempo en revisar los documentos producidos en la evaluación para someterlos a un juicio conjunto con esta. Como las dimensiones de los estándares deben depender mayoritariamente de la evaluación y sus capacidades, entonces este estándar, más que exigir a la contraparte su participación y entrega de información, debe asegurarse desde un principio que la contraparte tiene las capacidades y la disposición para realizar en conjunto un buen proceso evaluativo.

## **Conclusiones**

Finalmente, se pueden realizar conclusiones respecto a la importancia de la inclusión de los estándares a las prácticas evaluativas chilenas. A pesar de la generalidad y de ciertas falencias, es de gran importancia la incorporación de los estándares de evaluación a la práctica chilena porque establecen un marco de referencia para que los evaluadores auto observen sus propias prácticas en base al criterio de otros evaluadores expertos. Gracias a ello, pueden reconocer cuáles son sus fortalezas y cuáles son sus debilidades en la evaluación, permitiendo el control común de la rigurosidad que alcanzan estas. Además, la retroalimentación constante genera una comunidad científica que define las bases para integrar las investigaciones y evaluaciones que son parte del conocimiento acumulado por la profesión.

Por otra parte, los estándares permiten el diálogo en un mismo lenguaje entre los evaluadores nacionales. El análisis de la información presente en este artículo indica que la falencia en las prácticas comunicativas entre distintos actores lleva al gasto innecesario de tiempo, esfuerzo y recursos, especialmente de parte de evaluadores. Una comunicación fluida bajo los mismos parámetros podría ayudar a que las evaluaciones sean efectivas en compartir las experiencias de evaluación, superando errores presentes en varias de estas.

Víctor Salas fue enfático en mencionar que los estándares de evaluación presentaban falencias en la utilidad práctica que tenían para guiar una evaluación, debido a que eran muy generales. En específico, esto ocurre porque no se logran establecer parámetros que soporten la especificidad de las evaluaciones. Los estándares de evaluación propuestos por la ReLAC son demasiado completos, y por ello, no consideran que hay evaluaciones que no tienen la capacidad ni la necesidad de abordar todas sus dimensiones. Evaluaciones de programas, como las realizadas por el entrevistado, están focalizadas en aspectos que son altamente relevantes y que no se ajustan a los estándares propuestos. No todas las evaluaciones pueden ser orientadas por la suma de los estándares, sino, sólo por parte de estos. Por tanto, parece ser que una opción para mejorar los estándares de evaluación (que, como se vio, son importantes) es desagregarlos por tipo de evaluación que se realice, complejizando su esquema para que se adapte a las condiciones y posibilidades de cada evaluación. Esto le permitiría una flexibilidad útil en la práctica evaluativa.

Por ello, esta investigación debe entrar en diálogo con los estándares de evaluación de ReLAC y con otros que han sido propuestos por otras instituciones para generar avances en el área de la evaluación. La postura de expertos permite entender cuáles son las fortalezas y debilidades de los parámetros y cuáles son las vías para que estos puedan reunir bajo su alero a la mayor cantidad posible de evaluaciones y evaluadores.

## Bibliografía

- African Evaluation Association. 2007. *AfrEA Evaluation Guildlines*. Último acceso: 6 de 4 de 2017.
- American Evaluation Association. 2013. *American Evaluation Association Guiding Principles For Evaluators*. Último acceso: 11 de 4 de 2017. <http://www.eval.org/p/cm/ld/fid=51>.
- American Evaluation Association. 2004. «Guiding Principles for Evaluators.» *American Evaluation Association* 2-9.
- Australian Evaluation Society . 2013. «Guidelines for the ethical conduct of evaluations.» *Australasian Evaluation Society* 2-12.
- Australian Evaluation Society. 2013. *Evaluation Journal of Australasia*. Último acceso: 9 de 4 de 2017. <https://www.aes.asn.au/evaluation-journal-of-australasia.html>.
- Bilella, Pablo, Sergio Martinic, Sarah Desireé, Luis Soberón, Ana L Guzmán, and Esteban Tapella. 2016. *Estándares de evaluación para América Latina y el Caribe*. Buenos Aires: Red de Seguimiento, Evaluación y Sistematización de América Latina y el Caribe.
- Canadian Evaluation Society. 2011. *Program Evaluation Standards*. Último acceso: 8 de 4 de 2017. <https://evaluationcanada.ca/program-evaluation-standards>.
- Patel, Mahesh. 2013. «African Evaluation Guidelines.» *African Evaluation Journal* 1 (1): 1-5.
- The Canadian Evaluation Society. 2010. «Competencies for Canadian Evaluation Practice.» *Canadian Evaluation Society* 1-15.

## Anexos

### Pauta de entrevista:



UNIVERSIDAD DE CHILE

## Entrevistas Estándares de Evaluación - Actores

Curso Evaluación de Proyectos Sociales

Universidad de Chile

Agradecemos su participación en esta investigación y la firma de nuestro Consentimiento Informado para su realización, así como la autorización para grabar la conversación. Como le adelantamos en él, esta entrevista durará aproximadamente 1 hora y en ella le propondremos conversar sobre diversos temas o problemas relacionados la práctica de la evaluación.

Esta entrevista tiene el propósito de someter al juicio de expertos chilenos los estándares de evaluación recientemente publicados por la Red de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe (RELAC). En general, la intención de estos estándares es contribuir al desarrollo de un marco común de referencia que sirva de orientación para la obtención de evaluaciones que cumplan con cinco dimensiones, a saber: evaluación rigurosa, adecuada evaluabilidad, evaluación conducida de acuerdo con principios éticos y jurídicos, adecuada comprensión cultural, y finalmente relevancia y utilidad.

Para esto, le definiré brevemente cada una de las dimensiones y luego le haré preguntas específicas respecto a las subdimensiones que contiene. En específico, buscamos saber si estos estándares de evaluación ya están integrados en la práctica evaluativa actual, si cree pertinente que sean integrados, y qué dificultades conllevaría su incorporación. También, queremos conocer su propia experiencia profesional respecto a la integración de estos estándares.

Pauta de Entrevista.

Dimensión/presentación	Preguntas
<p>Dimensión: <b>Evaluación rigurosa.</b> La dimensión evaluación rigurosa permite dar cuenta de lo que se observa en términos de análisis, de manera precisa y exacta. Los estándares que corresponden a esta dimensión deben garantizar una evaluación estrechamente vinculada a su contexto, que sea válida y cuya metodología sea rigurosa en cuanto al diseño, planificación y orientación hacia una evaluación de calidad, por tanto, debe generar resultados precisos y evidenciar recomendaciones útiles para la toma de decisiones.</p>	<ul style="list-style-type: none"> <li>• <b>Pregunta inicial:</b> ¿Qué elementos cree ud que se deben considerar para realizar una evaluación de calidad?</li> <li>• <b>Contextualizar la evaluación:</b> ¿Usted ha caracterizado el contexto social, político, económico y cultural al momento de realizar intervenciones?, en caso de que la respuesta sea afirmativa ¿Cómo lo ha hecho? ¿Comprende dentro de la intervención factores externos que inciden en la intervención? ¿Usted ha abordado criterios que sitúen las intervenciones que ha realizado de acuerdo a la especificidad de cada localidad?</li> <li>• <b>Descripción detallada del objeto de evaluación:</b> En su experiencia ¿Se detallan las acciones, productos y propósitos del objeto de manera suficiente o siente que han faltado elementos para que lo sea? En caso de respuesta positiva ¿cómo lo han logrado? En caso de respuesta negativa ¿Por qué creen que no se ha logrado?</li> <li>• <b>Preguntas evaluativas relevantes:</b> En su experiencia ¿Cómo ha definido los propósitos y preguntas de evaluación? ¿Cree que ha sido pertinente para los principales actores de la evaluación? ¿Por qué?</li> <li>• <b>Metodología válida y confiable:</b> ¿Bajo qué criterios ha valorado el diseño de la evaluación? ¿Qué elementos han dificultado y facilitado la construcción del diseño?</li> <li>• <b>Niveles de participación adecuados de actores involucrados:</b> En su experiencia ¿Se han abierto espacios de participación con los actores involucrados? ¿Se ha limitado la participación de algún actor? ¿Por qué?</li> <li>• <b>Conclusiones relevantes:</b> ¿Qué elementos son fundamentales para la elaboración de las conclusiones de la evaluación? ¿Cómo ha vinculado usted las conclusiones a los propósitos de la evaluación?</li> <li>• <b>Recomendaciones útiles y factibles:</b></li> <li>• ¿En base a qué criterios ha formulado las recomendaciones a la contraparte? ¿Para formular las recomendaciones, ha considerado a la población objetivo? ¿De qué manera?</li> <li>• <b>Reportes y comunicación pública efectiva:</b> ¿Cuáles han sido los alcances y limitaciones para presentar los resultados de la evaluación? ¿Con qué actores considera importante comunicar los resultados de la</li> </ul>

<p><b>Dimensión: Adecuada evaluabilidad.</b> La dimensión de Adecuada evaluabilidad tiene por objeto garantizar que la evaluación sea realista, prudente, apropiada, y consciente de los costos que implica. Por ello, antes del desarrollo de la evaluación, debe definirse adecuadamente el objeto a evaluar y si la metodología permite obtener evidencia que informe sobre la hipótesis y las preguntas de evaluación planteadas. Así, los estándares de esta dimensión procuran determinar si vale la pena realizar la evaluación en las condiciones que se ha establecido.</p>	<p>evaluación? ¿Por qué?</p> <ul style="list-style-type: none"> <li>• Pregunta inicial: ¿Qué importancia le otorga a una adecuada evaluabilidad de la evaluación? ¿Cuán importante cree que es establecer condiciones previamente?</li> <li>• <i>Sobre gestión efectiva de la evaluación.</i> ¿Cómo se construyen estrategias de gestión que guíen una evaluación exitosa? ¿Cuál es la flexibilidad adecuada de estas? Pensando en algunas evaluaciones en las que ha participado ¿Cómo se ha aplicado?</li> <li>• <i>Procedimientos prácticos.</i> ¿Cómo se adecúan los procedimientos de evaluación al objeto de estudio?</li> <li>• <i>Viabilidad contextual, social y política</i> ¿Cree que las evaluaciones deben considerar los intereses políticos y culturales específicos que enfrenta? ¿O más bien debería realizarse un procedimiento estándar generalizado?</li> <li>• <i>Realista</i> ¿De qué manera se puede lograr una coherencia entre recursos financieros, humanos y temporales para el desarrollo de la evaluación? <i>Refiérase a su experiencia</i></li> </ul>
<p><b>Dimensión: Evaluación conducida de acuerdo con principios éticos y jurídicos.</b> La dimensión de “Evaluación conducida de acuerdo a principios éticos y jurídicos” refiere a que el proceso sea transparente y legítimo, garantizando que los derechos de todos los actores involucrados sean respetados. Esto se logra a través de procedimientos que no afecten la integridad de las personas, ya sean parte de una institución o una comunidad.</p>	<ul style="list-style-type: none"> <li>• Pregunta inicial: ¿Desde su opinión, qué rol cumplen los principios éticos y jurídicos en el proceso de evaluación?</li> <li>• <i>Respeto del derecho de las personas</i> ¿Qué derechos tiene la persona que otorga información al evaluador? ¿Qué limitaciones pueden significar estos derechos durante el proceso de evaluación?</li> <li>• <i>Autonomía</i> En su experiencia, ¿Ha sido posible mantener la función de evaluación independiente de otras funciones de gestión? ¿Qué medidas aseguran la imparcialidad con la contraparte?</li> <li>• <i>Transparencia</i> ¿Cuál es su opinión sobre hacer públicos los resultados de una evaluación? ¿Qué limitaciones y beneficios tiene establecer criterios de confidencialidad con la contraparte?</li> <li>• <i>Legalidad</i></li> </ul>

	<p>¿Cuáles son los elementos más relevantes a los que se debe llegar a acuerdo con la contraparte antes de comenzar la evaluación? ¿Qué grado de flexibilidad deberían tener estos acuerdos?</p>
<p>Dimensión: <b>Adecuada comprensión cultural.</b> La dimensión Adecuada comprensión cultural alude a que las evaluaciones deben respetar lenguajes, códigos y sentidos culturales de las distintas poblaciones afectadas o beneficiadas por las evaluaciones. Estos procesos de evaluación deben ser sensibles (en el sentido de adaptarse) a diversos contextos históricos y culturales, con una buena comunicación y empatía.</p>	<ul style="list-style-type: none"> <li>• Pregunta general para empezar el tema: ¿Cree usted que las evaluaciones deben ser sensibles a diversos contextos históricos y culturales en las cuales se presentan?</li> <li>• <i>Sobre igualdad y equidad:</i> Dentro de esto ¿Qué potencialidades y dificultades cree que implica tratar a cada persona sin importar su condición? (jerarquía, condición social, económica y/o cultura)</li> <li>• <i>Sobre derechos culturales:</i> ¿Piensa que deba existir un protocolo para tratar temas sensibles para una comunidad? ¿Qué desventajas puede traer utilizar un protocolo para cuidar el perjuicio de la comunidad?</li> <li>• <i>Sobre reciprocidad:</i> ¿De qué manera cree que puede lograrse un respeto por valores, modos de pensar y conocimiento de las comunidades que no sea en desmedro de la evaluación? Y ¿Cuál ha sido su experiencia al respecto?</li> </ul>
<p>Dimensión <b>Relevancia y utilidad:</b></p> <p>La dimensión de relevancia y utilidad busca garantizar una mayor eficiencia en las evaluaciones para que sirvan a las necesidades de información y conocimiento de los usuarios. Por eso, deben ser relevantes, oportunas y prácticas para aportar de manera eficiente a la toma de decisiones.</p> <p>Los estándares propuestos pretenden incrementar la probabilidad de que los participantes valoren los procesos y productos de la evaluación, encontrándolos significativos para satisfacer necesidades. De este modo, hay interés en lo relevante y práctico para sugerir recomendaciones útiles para la toma de decisiones y el mejoramiento del programa, contribuyendo al desarrollo del</p>	<ul style="list-style-type: none"> <li>• Pregunta inicial: ¿A su juicio, qué elementos debe considerar una investigación evaluativa para ser relevante y útil?</li> <li>• <i>Participación efectiva</i> ¿De qué manera ha incorporado las perspectivas de diversos actores involucrados e interesados en sus investigaciones evaluativas? ¿cómo cree Ud. que se debe promover la participación efectiva de dichos actores?</li> <li>• <i>Propósitos acordados mutuamente</i> ¿De qué forma son acordados los propósitos? ¿Cree Ud. que han sido claros y precisos en sus investigaciones?</li> <li>• <i>Valores explícitos</i> ¿Qué papel juegan sus propios juicios y puntos de vista en el proceso evaluativo?</li> <li>• <i>Información relevante, pertinente y oportuna</i> ¿Considera Ud. que la información producida en sus investigaciones ha resultado de utilidad para las</li> </ul>

<p>conocimiento para su uso en diferentes contextos.</p>	<p>necesidades de los grupos interesados? ¿De qué manera?</p> <ul style="list-style-type: none"> <li>• <i>Resultados útiles</i> ¿En qué medida considera Ud. que los resultados de sus investigaciones evaluativas han permitido a los participantes redescubrir, reinterpretar o revisar sus percepciones y comportamientos sobre el objeto a investigar?</li> <li>• <i>Comunicación y reportes puntuales y apropiados</i> ¿Qué importancia le da Ud. a la comunicación con la contraparte durante todo el proceso evaluativo?</li> <li>• <i>Interés por las consecuencias e incidencia</i> ¿De qué forma las evaluaciones y un buen uso de sus resultados pueden incidir en el mejoramiento de la calidad de vida? ¿Hasta qué punto las evaluaciones deberían intervenir en la vida de los actores involucrados? De qué forma son acordados los propósitos de la evaluación? ¿Cree Ud. que han sido claros y precisos para todos los actores involucrados en sus investigaciones? ¿Por qué?</li> <li>• <i>Valores expuestos</i> ¿Cómo se hace cargo de sus propios juicios y puntos de vista en el proceso evaluativo? ¿Los incluye, excluye, los manifiesta?</li> </ul>
--	--

Antes de concluir ¿quisiera agregar algo más o tiene alguna reflexión? Finalmente le agradecemos el tiempo que nos ha entregado. Le haremos llegar una transcripción anonimizada de esta reunión y también cualquier resultado de la investigación correspondiente.