

Coloquio. La evaluación y los Objetivos de Desarrollo Sostenible (ODS)

Evaluation and SDGs. Colloquium with María Bustelo, Osvaldo Feinstein y Oscar García

Con María Bustelo, Osvaldo Feinstein y Oscar García

Como es habitual en Cuadernos del CLAEH, el coloquio propone un recorrido por una temática de interés que el CLAEH entiende relevante y central, no solo por su actualidad o importancia coyuntural, sino también porque permite una mirada multidisciplinaria y plural sobre fenómenos nacionales, regionales e internacionales. En este caso, el 22 de octubre de 2020, María Bustelo,¹ Osvaldo Feinstein,² y Oscar García³ conversaron

- 1 Doctora en Ciencias Políticas y Sociología (Universidad Complutense de Madrid, UCM). Máster en Evaluación de Programas (Universidad de Illinois, Estados Unidos). Profesora de Ciencia Política y de Administración en la UCM. Experta en evaluación de políticas públicas y directora del Máster en Evaluación de Programas y Políticas Públicas (UCM) desde sus inicios en 2002. Es autora de diversas publicaciones sobre evaluación de políticas públicas, políticas de género en España y en Europa, y *mainstreaming* de género. Formó parte del Global Evaluation Advisory Committee (GEAC) de ONU Mujeres (2014-2019). Es coordinadora europea del proyecto Horizonte 2020 SUPERA (Supporting the Promotion of Equality in Research and Academia 2018-2022), y ha sido la investigadora responsable en la UCM de otros proyectos europeos de investigación sobre políticas de género en España y en Europa, así como de proyectos nacionales I+D, el último de los cuales UNIGUAL (Políticas de Igualdad en la Universidad Española, 2018-2021). Integra los comités editoriales de las revistas *Evaluation*, *Evaluation Program & Planning*, *European Journal of Politics and Gender*.
- 2 Profesor en el Máster en Evaluación de Programas y Políticas Públicas de la Universidad Complutense de Madrid. Asesor de la Oficina Independiente de Evaluación del Global Environmental Facility (GEF). Miembro del Panel Internacional Asesor de la Oficina de Evaluación Independiente del PNUD. Ha sido profesor en universidades e instituciones europeas y latinoamericanas. Fue asesor y gerente en el Departamento de Evaluación del Banco Mundial y miembro del US National Academy of Sciences Panel on Evidence Based Policy in the Social Sciences y del Panel de Seguimiento y Evaluación del Consejo Científico del Grupo Consultativo para la Investigación Agrícola Internacional (CGIAR); evaluador principal del FIDA y consultor del Centro Internacional de Investigaciones del Desarrollo (IDRC), de la Agencia de Evaluación y Calidad (en España) y de organismos internacionales como Grupo de Evaluación de las Naciones Unidas, FAO, Fondo de Ciencia y Técnica de Naciones Unidas, UNESCO, Departamento de Cooperación Técnica de la ONU, Caribbean Development Bank, Banco Interamericano de Desarrollo, Banco Africano de Desarrollo, CLEAR AA, Banque de France, OIT, CEPAL e ILPES. Es autor y editor de libros y artículos sobre evaluación y desarrollo.
- 3 Economista. Máster en Gestión del Cambio Organizacional por la New School University (Nueva York, Estados Unidos) y MBA de la Universidad Católica Boliviana en asociación con el Instituto de Desarrollo Internacional de Harvard. Es director de la Oficina de Evaluación Independiente del PNUD, oficina que proporciona dirección estratégica, liderazgo intelectual y evaluaciones con credibilidad para avanzar en la labor del PNUD de ayudar a los países en su progreso hacia la consecución de los ODS nacionales.

vía Zoom® con los editores asociados a este número, Leopoldo Font y Juan Pablo Mótola, tomando como referencia el tema central del dossier «Evaluación y desarrollo sostenible».

Sobre el campo de la evaluación de proyectos, programas y políticas públicas: avances y desafíos

LEOPOLDO FONT (LF) —Como mencionábamos en la convocatoria, y nos gustaría volver a compartir con ustedes para centrar el tema en los aspectos que con Juan Pablo quisimos destacar para este dossier, los desafíos crecientes que afrontan los Estados democráticos para alcanzar el desarrollo sostenible hacen necesario el fortalecimiento de políticas públicas orientadas a prioridades, objetivos y metas que conjuguen de forma integral e indivisible las dimensiones socioculturales, económicas, políticas, institucionales y ambientales. La complejidad y la profesionalización de la gestión pública ha intensificado la necesidad de contar con procesos de toma de decisión basados en evidencias confiables, válidas y oportunas. Esencialmente para optimizar su desempeño y apropiación por la ciudadanía, en un entorno además rodeado de incertidumbres, ni qué hablar ahora en este contexto de pandemia, y con la necesidad de contar con una oportuna rendición de cuentas. Del mismo modo, estas exigencias requieren la generación de dispositivos institucionales que permitan facilitar el aprendizaje organizacional, la rendición de cuentas y la participación de la ciudadanía en los asuntos públicos. En este marco cobra especial relevancia conocer y difundir experiencias de fortalecimiento e innovación vinculadas al monitoreo y la evaluación de proyectos, programas y políticas que se han llevado adelante en los últimos años.

Sugerimos empezar por un tema de aspecto bien general. En las últimas décadas, el campo de la evaluación de proyectos, programas y políticas públicas ha vivido una evolución muy importante. Como primeras preguntas: ¿cuáles creen que son los mayores logros alcanzados en ese campo?, ¿cuáles son los principales aspectos que aún es necesario desarrollar en el campo de la evaluación para potenciar su efectiva utilización?

Tiene más de 25 años de experiencia en la dirección de áreas de planificación estratégica, evaluación, gestión basada en resultados, agricultura, desarrollo rural, medio ambiente, gestión de recursos naturales, reducción de la pobreza y desarrollo humano. Fue director de la Oficina Independiente de Evaluación en FIDA y jefe de los Servicios de Asesoramiento para la Economía Verde del PNUMA. Es autor de libros y artículos sobre evaluación del desarrollo.

OSVALDO FEINSTEIN (OF) — A mí me ha tocado estar desde hace muchos años en estos temas. Comparando el estado de la evaluación en los ochenta con el actual, por entonces había muy poca gente trabajando en estos temas en América Latina. Es el caso de un colega y amigo uruguayo, Humberto *Chacho* Costa,⁴ que tuvo en Paraguay un papel importante en la institucionalización de la evaluación, uno de los avances que se dio en las décadas siguientes. Este fue uno de los grandes logros. En los ochenta, en la mayor parte de los países de América Latina no había prácticamente nada en términos de instituciones dedicadas a la evaluación y al monitoreo. Hoy en día son pocos los países donde no hay alguna institución que tenga este tipo de responsabilidad.

En los ochenta, en la mayor parte de los países de América Latina no había prácticamente nada en términos de instituciones dedicadas a la evaluación y al monitoreo. Hoy en día son pocos los países donde no hay alguna institución que tenga este tipo de responsabilidad. OSVALDO FEINSTEIN

Segundo, en términos de capacidades, en los ochenta prácticamente no había ningún programa, y en los noventa había un solo programa en América Latina, incluso en el mundo, que era el PREVAL, apoyado por el FIDA.⁵ Hoy en día ya hay varios programas, y varios que operan en América Latina. Un tercer aspecto relacionado con este es que, en esa época —incluso en los noventa—, había muy poca demanda de evaluación, tanto del sector público como de la sociedad civil. Más bien parecía una disciplina inexistente; era muy difícil armar un espacio para la discusión. Finalmente, creo que vale la pena también mencionar el avance en métodos, en técnicas, relacionados con los otros aspectos.

La pregunta tenía una segunda parte sobre lo que queda por hacer y sobre todo desde el punto de vista de la utilización de las evaluaciones. Al fin y al cabo, sabemos que las evaluaciones tienen costos y que los beneficios dependen de la utilización, entonces es muy importante ver cómo podemos avanzar en esa línea. Hay dos áreas que son desafíos, en las que se ha avanzado también pero queda mucho por hacer. Una es la

4 Humberto Acosta Fernández, uruguayo, ingeniero agrónomo especialista en proyectos de inversión. Reside actualmente en Asunción, Paraguay. Se desempeñó como consultor de organismos financieros multilaterales y de asistencia técnica en diez países de América Latina.

5 El Programa para el Fortalecimiento de la Capacidad Regional de Seguimiento y Evaluación para la Reducción de la Pobreza Rural en América Latina y el Caribe (PREVAL), financiado por el Fondo Internacional de Desarrollo Agrícola (FIDA), implementado a partir de 1997, busca desarrollar las capacidades locales para el diseño y desarrollo de sistemas de seguimiento y evaluación participativos orientados al impacto.

comunicación, que requiere no solo que los evaluadores comuniquen, sino también que haya una división de trabajo involucrando gente con competencia, especialmente en comunicación —inclusive periodistas, escritores—, que pueden apoyar mucho el trabajo de los evaluadores para que la evaluación no sea un mero documento. Y relacionado también con este punto, todo lo que tiene que ver con la *oportunidad* de las evaluaciones, la agenda, para que cuando se termine la evaluación se la pueda utilizar y, cuando hay decisiones importantes que tomar, que haya posibilidad de recurrir a evaluaciones que aporten evidencias.

LF —En las décadas de los ochenta y noventa los que pedían evaluaciones de proyectos al CLAEH y a las ONG en Uruguay eran los organismos de cooperación internacional.

OF — Es importante lo que mencionas porque en realidad en los ochenta y buena parte de los noventa la evaluación era un requisito, algo que se exigía. Esa demanda no era algo que surgía de la sociedad civil, ni siquiera de los gobiernos; se hacía para cumplir con un requisito, un ritual. Eso cambió.

OSCAR GARCÍA (OG) —Osvaldo da una perspectiva histórica importantísima sobre el papel de la evaluación y cómo ha evolucionado en la región. Creo que, adicionalmente, hoy en día existe una mayor conciencia sobre la importancia de la evidencia para mejorar la calidad de la toma de decisiones, particularmente de la formulación de políticas públicas. Hay una participación del mundo académico que está empujando por una agenda de política pública basada en evidencia que es fundamental en los días de hoy. Como ustedes han podido observar, la crisis del covid-19 ha puesto de manifiesto aún más el desprecio que algunas veces existe por la evidencia científica, y los grandes niveles de desinformación que se están dando en la sociedad a partir de interpretaciones ideologizadas a un lado o a otro lado de determinado fenómeno social. Por eso creo que esta mayor cantidad de personas dedicadas al campo de la evaluación es el resultado de una mayor conciencia sobre la importancia de la evidencia.

Hoy en día existe una mayor conciencia sobre la importancia de la evidencia para mejorar la calidad de la toma de decisiones, particularmente de la formulación de políticas públicas. OSCAR GARCÍA

Un segundo aspecto, reforzando lo que decía Osvaldo con relación al avance metodológico, es que ha existido un importante reconocimiento de la evaluación de impacto. Con todos los métodos contrafactuales haciendo la pregunta «¿qué hubiese suce-

dido si esta intervención no tenía efecto?», la evaluación de impacto ha tenido últimamente un reconocimiento. Muchos pueden cuestionar la validez de este reconocimiento, pero incluso a los premios nobel de Economía de 2019, Abhijit Banerjee,⁶ Esther Duflo⁷ y Michael Kremer,⁸ les dan un reconocimiento por ver qué funciona y qué no funciona. Como ustedes saben, en el campo de la evaluación hay distintas posturas y hay quienes ven la limitación de estos enfoques de las aproximaciones basadas en los *randomized controlled trials*⁹ para determinar la real complejidad de un proceso de desarrollo; eso es evidente y yo estoy en esa escuela, pero no podemos dejar de reconocer que en términos de los métodos de validación hay un préstamo de instrumental metodológico de las ciencias naturales, particularmente de la medicina, que se está aplicado ahora, en poca escala, a las ciencias sociales.

Creo que ha sido un avance importante porque este lenguaje de la medición de impacto ha llegado también a la opinión pública, que está hablando más del impacto de estas medidas en la calidad de vida de las personas. Ese es un tema muy importante para la agenda de desarrollo sostenible: ¿cómo podemos medir el impacto de las políticas públicas en una agenda de sostenibilidad, que tenga en cuenta los cuatro aspectos de la sostenibilidad: el aspecto económico, el social, el ambiental y la gobernabilidad?

Finalmente, ratificando también lo que dice Osvaldo, creo que hay una mayor diseminación y proliferación de las asociaciones profesionales de evaluación. Hoy en día estas asociaciones juntan hombres y mujeres de distintas formaciones académicas para tratar de avanzar en una agenda de la importancia de la evaluación.

Ahora bien, ¿cuáles son las limitaciones o los desafíos? Yo diría, en primer lugar, que juntar la evidencia no es suficiente. Necesitamos marcos interpretativos y, en estos marcos interpretativos, una mayor aproximación con la teoría y con la ciencia. Simplemente juntar evidencia no es un trabajo de evaluación. Para preguntar qué funciona y qué no funciona, y particularmente para tratar de responder a la pregunta de «¿por qué las cosas no funcionan?» necesitamos afincar la evaluación en teoría. Y esto creo que es todavía un desafío.

6 Economista indio-estadounidense. Profesor internacional de Economía de la Ford Foundation en el Massachusetts Institute of Technology (MIT).

7 Economista francesa, cofundadora y directora del Abdul Latif Jameel Poverty Action Lab (J-PAL), y profesora de Reducción de la Pobreza y Economía del Desarrollo en el MIT.

8 Economista estadounidense. Profesor de la Universidad de Chicago.

9 Una prueba controlada aleatorizada o ensayo controlado aleatorizado (en inglés, *randomized controlled trial*, RCT) es un tipo de experimento científico usado principalmente para conocer el efecto de una intervención gubernamental sobre una población. Se considera la forma más fiable de evidencia científica porque si se implementa correctamente puede eliminar todos los posibles sesgos en la estimación del efecto de la intervención. Los RCT se utilizan principalmente en investigación clínica y en ciencias sociales.

Necesitamos transitar hacia modelos y sistemas de evaluación más sistémicos, que puedan dar cuenta de problemas sociales más complejos que aquejan a nuestra sociedad. Temas como el cambio climático, el incremento de las desigualdades en nuestras sociedades no se pueden interpretar exclusivamente con una lógica lineal de causa y efecto. Hay múltiples componentes, y esa multiplicidad tiene que ser reconocida por estos esquemas interpretativos. OSCAR GARCÍA

En segundo lugar, muchos de los esquemas interpretativos usados actualmente son lineales en su lógica, en una relación causa y efecto: esta es la causa, este es el efecto; si yo entiendo la causa, cambio el efecto. Y sabemos que necesitamos transitar hacia modelos y sistemas de evaluación más sistémicos, que puedan dar cuenta de problemas sociales más complejos que aquejan a nuestra sociedad. Temas como el cambio climático, el incremento de las desigualdades en nuestras sociedades no se pueden interpretar exclusivamente con una lógica lineal de causa y efecto. Hay múltiples componentes, y esa multiplicidad tiene que ser reconocida por estos esquemas interpretativos. Creo que la evaluación todavía tiene un camino a recorrer en incorporar estas miradas más sistémicas.

Finalmente, el tercer aspecto es un mejor uso sobre las tecnologías de comunicación aplicadas al campo de la evaluación, tanto para la recolección de datos como para la interpretación y la diseminación de estos. Vemos que el diálogo habita sobre los datos; esto es muy importante pero está siendo dominado por los generadores de datos, no necesariamente por quienes pueden ser usuarios muy importantes de estos datos, como son los evaluadores, para informar cómo se mejora la calidad de la toma de decisiones.

LF —En cuanto al tema de la evaluación de impacto, dejo anotado para seguir trabajando el tema de la mitología de lo no posible, de lo caro, de todo lo que hay que avanzar para romper obstáculos sobre todo epistemológicos que nos generamos.

MARÍA BUSTELO (MB) —Tenía apuntadas cosas un poco más generales pero las voy a enganchar con otras cuestiones que han comentado Osvaldo y Oscar. Coincido con Oscar en que lo primero que me ha salido cuando he pensado en los mayores logros es justamente el reconocimiento y el respeto a la diversidad metodológica y a los temas y los asuntos más transversales y más difíciles de aprehender. El reconocimiento y respeto a la diversidad metodológica es uno de los grandes logros del campo de la evaluación. Oscar resaltaba la importancia de los últimos desarrollos de evaluaciones de impacto que ha habido, no sólo con los *randomized controlled trials* sino incluso con otro



María Bustelo

tipo de búsquedas, a lo mejor, más en el ámbito de la contribución y no tanto de la atribución, y cómo estamos explorando eso. Creo que es uno de los campos en los que más se ha fomentado la diversidad metodológica y que más se ha reconocido; de hecho, creo que el campo de la evaluación es el que ha ayudado más a hablar al concepto de métodos mixtos y su utilización. Y lo digo en una doble dirección. Recuerdo que una vez alguien del ámbito europeo me dijo: «eso de los métodos mixtos es lo que dice la gente cualitativa cuando no se pueden utilizar los datos cuantitativos». Realmente lo digo con toda la razón de ser, o sea, el respeto y el reconocimiento a la diversidad metodológica, de métodos que no quiero llamarles alternativos —porque al decirle *método alternativo* estás reconociendo que hay alguno que es el *golden standard*—. Creo que esa variedad metodológica es importante y hay que cultivarla, porque es lo que consigue acertar en diferentes niveles, a diferentes sectores, en el tema de la evaluación.

Por supuesto, añadido el que desarrollo de una comunidad profesional evaluadora con un carácter muy internacional, que hemos visto crecer en estas últimas décadas, es uno de los grandes valores añadidos de la evaluación. Y yo diría, también, la acumulación de un cuerpo de literatura y de un bagaje teórico reconocido e identificado por

quienes hacemos evaluación, lo que, unido a lo anterior, añade también una cierta identificación profesional. Por supuesto, el desarrollo formal de la formación académica de una manera un poco más adecuada, con todo lo que conlleva: que haya cursos de posgrado, que haya tesis, que se defiendan tesis en el tema de evaluación, lo que, por supuesto, ha sido en paralelo también con un desarrollo organizativo en diferentes lugares, probablemente, mucho más en organismos de carácter internacional que en organismos de tipo nacional muy variado, dependiendo de los países. Como bien decía Osvaldo, es algo que en estos años hemos visto crecer, que llevaría a un cierto grado de institucionalización.

Respecto al grado de institucionalización es muy interesante ver el trabajo que han empezado a hacer Reinhard Stockmann y Wolfgang Meyer, la gente que está trabajando en la Universidad del Sarre, Alemania, y que también trabajan mucho en Costa Rica, que han hecho un libro que examina el progreso de la institucionalización de la evaluación en los países europeos desde varias perspectivas.¹⁰ El primer volumen es el de Europa, salió en abril, en mitad de la pandemia, y está a punto de publicarse el de las Américas. Me llevé con esto una de las mejores alegrías. El 30 de setiembre compartí un webinar del International Program for Development Evaluation Training (IPDET) con Stockmann, Meyer y Claudia Maldonado,¹¹ y estábamos hablando sobre ello, y curiosamente decía Stockmann que, comparando con Europa lo que han visto del volumen de las Américas, sobre todo de Latinoamérica, se notaba que, a pesar de que había por ejemplo carácter normativo, leyes, incluso estructuras, en mayor cantidad de lo que puede haber a veces en Europa, sin embargo, había una brecha importante y masiva respecto a la implementación. Lo que quería decir era que, luego, prácticamente no había un uso de esas evaluaciones en sistemas de evaluación. No se trata de saber solo cuántas estructuras hay y cómo se han generado —incluso lo de la demanda, que comentaba Osvaldo—, sino también, y sobre todo, si esto se está realmente utilizando, si está teniendo una incidencia real.

Yo me he puesto un poco más poética pensando en el futuro. Porque lo que decía es que la evaluación nos tiene que ayudar sobre todo a la comprensión de este loco mundo que estamos viendo. Creo que estamos en un buen momento para hacerlo; pero es no solo reconocimiento de la complejidad, de —como bien ha dicho Oscar— la sostenibilidad y diferentes aspectos, el reconocimiento cada vez mayor de que hay muchas desigualdades de carácter estructural y que son las que se nos rebelan, por mucho que hagamos, y que la evaluación nos tiene que ayudar a saber cómo realmente operan, a

10 Stockmann, R., Meyer, W., y Taube L. (eds.). (2020). *The Institutionalisation of Evaluation in Europe*. Londres: Palgrave Macmillan.

11 *The Evaluation Globe - Has Evaluation Failed? With R. Stockmann, W. Meyer, C. Maldonado, M. Bustelo*. IPDET Mini Series n.º 4, Disponible en <https://ipdet.org/mini-series-4>.

saber cómo combatir las mejor. Estarían los viejos problemas y también los nuevos problemas que surgen, y cada vez tenemos que estar más preparados para ello. ¿Quién nos iba a decir, hace un año nada más, que íbamos a estar así, que nuestra máxima preocupación sería algo que no nos imaginábamos de ninguna manera? Es decir, la evaluación debe ayudar no solo a explicar lo que funciona y lo que no, sino a comprender mejor. Es una cuestión de profundidad, es comprender mejor los problemas públicos que ahora tenemos y que seguirán viniendo, y los nuevos problemas. Para eso también hay que pensar metodológicamente cómo hacerlo; por tanto, debemos saber cómo responder a la complejidad y a la incertidumbre, reconociendo a la primera sin simplificarla en exceso (porque entonces dejamos de comprenderla) y tratando de reducir la segunda, probablemente no siendo demasiado ambiciosos al tratar de reducir toda la incertidumbre, pero al menos intentando reducirla de alguna manera.

La evaluación debe ayudar no solo a explicar lo que funciona y lo que no, sino a comprender mejor. Es una cuestión de profundidad, es comprender mejor los problemas públicos que ahora tenemos y que seguirán viniendo, y los nuevos problemas. Para eso también hay que pensar metodológicamente cómo hacerlo.

MARÍA BUSTELO

Me parece clave también que consigamos transmitir la idea y que exista una demanda de evaluación desde la sociedad civil. Se ha visto que esto prácticamente no opera en la institucionalización, ni siquiera en Europa. Yo quedé muy sorprendida porque pensé que la falta de demanda y de un discurso de evaluación en la sociedad era algo que pasaba más bien en España, pero vi que todos los colegas que habían escrito sobre los diferentes países europeos tenían exactamente la misma conclusión. Creo que esto es importante, tiene que haber un conocimiento, tiene que haber un discurso de evaluación por parte de la sociedad.

Y a esto lo relaciono con la reflexión de Oscar. Creo que en este momento tenemos que saber vender la evaluación como la necesidad de tener evidencia para contrarrestar y combatir las *fake news*, así de claro. Tenemos que generar una cultura de creencia, de empeño, en la evidencia. Otra de las cosas que considero interesantes como parte de ese bagaje teórico que se ha ido haciendo en estos años es el concepto de *pensamiento evaluativo*, que también tiene que ver con lo que decía Oscar de que no basta con recoger evidencia, que no es eso, hay más. Y en esa idea de pensamiento evaluativo está la idea de combatir las *fake news*, Creo que esto va a ser muy impactante.

Debemos afianzar la identidad de la evaluación y su valor añadido como instrumento para profundizar y defender la democracia, los valores democráticos. En ese sentido, creo que es cierto lo que comentaba Osvaldo de la comunicación, creo que es absolutamente clave, junto con la idea de que eso tiene que contribuir a ampliar el concepto de *uso* —que incluso Carol Weiss amplió de muchas maneras— a la idea de incidencia y de incidencia política. Esto lo que hay que conseguir de las evaluaciones.

Debemos afianzar la identidad de la evaluación y su valor añadido como instrumento para profundizar y defender la democracia. MARÍA BUSTELO

Se trataría de involucrar realmente a los destinatarios de la evaluación, sean quienes sean y a quienes hayamos identificado como tales. Los sistemas de evaluación se deberían volcar mucho más claramente hacia el seguimiento de ese uso y de esa incidencia. Deberíamos hacer, en la medida de lo posible, historias de éxito en el uso y en la incidencia, no solo en el uso instrumental de la información para cambiar cosas en la política evaluada, sino incluso más allá del aprendizaje organizativo, es decir, de la idea de que las organizaciones que evalúan aprenden; ir un poco más allá, desde el uso conceptual y los posibles cambios de paradigma, es decir, de influencia académica de la evaluación, al uso y el impacto que tiene realmente en las personas beneficiarias de las políticas evaluadas. Me estaba acordando de algo que se hizo hace unos años, que partió de las sociedades de evaluación —la Sociedad Europea estuvo en su momento en ello y también en Latinoamérica, Esteban Tapella¹² y más gente—, con la idea de *evaluation that makes a difference*. Es decir, realmente qué es lo que se ha conseguido con las evaluaciones, no solo en las organizaciones que tienen que utilizar la evaluación, sino incluso qué es lo que ha cambiado en la vida de las personas a las que van dirigidas las políticas.

12 Argentino. Licenciado en Trabajo Social (Universidad Nacional de San Juan, Argentina). Máster in Development Studies (Institute of Social Studies, Holanda). Doctor en Ciencias Agronómicas, con especialización en Ecología Humana (Universidad Nacional de Córdoba, Argentina). Director del Programa de Estudios del Trabajo, el Ambiente y la Sociedad (<https://petas-unsj.org/>) de la Universidad Nacional de San Juan. Consultor independiente en temas de evaluación y fotógrafo documental. Docente e investigador en el Instituto de Investigaciones Socioeconómicas en la Universidad Nacional de San Juan.

Sobre la evaluación y su aporte a la institucionalidad democrática en el marco de los ODS

LF —A mí lo que me queda picando como desafío es la tensión entre la inmediatez de la noticia, de la necesidad del dato, del resultado y todo a lo que nos lleva el uso de las TIC, cómo aparecen las *fake news* y demás, y lo lento que a veces son los procesos de evaluación. ¿Qué tendríamos que hacer para encontrar de manera más rápida los resultados necesarios para dar respuesta a lo que planteaban de la calidad democrática, que es consecuencia del uso de la evaluación y de los resultados de la evaluación?

OF —Me interesa recoger también algo que planteaba María. Al inicio enfatizabas mucho el aporte de la evaluación a la comprensión, y cuando decías eso yo estaba pensando también en la transformación, el papel de la evaluación en la transformación, que fue lo que nos llevó a plantear hace un tiempo la evaluación dinámica. Pero lo que decías sobre la incidencia, la influencia, tiene que ver justamente con ese papel que puede tener la evaluación para la transformación. Esto se relaciona también con el fortalecimiento de la democracia y, en particular, de la democracia deliberativa. Hay un uso potencial muy importante que está muy lejos del uso real: el uso de la evaluación en los parlamentos. De una forma muy práctica, en los parlamentos hay comisiones asesoras, hay presupuesto para las comisiones asesoras, pero es excepcional que usen o requieran evaluaciones. De la experiencia chilena se ha hablado mucho, pero muy poco sobre un aspecto muy interesante, que es la discusión de la agenda de evaluación en el Parlamento para ver cuáles son los temas que interesan, o la discusión parlamentaria que va a haber. Esta consulta al parlamento permite identificar temas importantes para las evaluaciones y las comisiones asesoras podrían aprovechar esa oportunidad para encargar evaluaciones que generen evidencias para enriquecer la discusión parlamentaria. Ahí hay toda una serie de oportunidades. Se han hecho algunos avances en Asia pero en América Latina es una asignatura pendiente. Cabe mencionar que, en Estados Unidos, la entidad más importante en evaluación, la General Accountability Office (GAO),¹³ no es del Poder Ejecutivo, sino que apoya al Congreso. Tratar de replicar el modelo del GAO, como se ha intentado en varios países donde el parlamento no tiene un papel importante, es inadecuado, pero es interesante tomar en cuenta este tipo de experiencia.

Me interesó mucho también lo que planteaba Oscar sobre un enfoque sistémico, considerando el desafío de la complejidad. Esto está relacionado también con el segundo tema planteado, en relación con los desafíos para el desarrollo sostenible y la Agenda 2030. Hay un riesgo que podemos llamar de siloísmo. Silo fue un personaje muy singular

13 Véase <https://www.gao.gov>.

en Argentina, tenía toda una escuela. Era un predicador.¹⁴ Se hablaba de siloísmo para referirse a este movimiento. Yo me refiero en un sentido mucho más terrenal, el enfoque de los silos, abordando cada uno de los los diecisiete Objetivos de Desarrollo Sostenible (ODS) en forma independiente, cada uno por su cuenta. Este es un riesgo grande, porque los Objetivos están muy interconectados, y es muy importante tener esa visión sistémica a la que se refería antes Oscar.

Esto se conecta también con un riesgo que hay tanto en relación con la Agenda 2030 como con lo que planteaba Leo sobre qué hacer para tener informaciones más rápidas en la evaluación. El riesgo es la burocratización de la evaluación, ver a la evaluación solamente como la producción de informes, informes que requieren un largo proceso de más de un año para llegar a tener la ilusión del informe perfecto. Es imperfecto ya de por sí, porque va a llegar muy tarde, entonces se pierde la oportunidad. Y hay otras formas a través de las cuales se pueden comunicar resultados de las evaluaciones. Hoy en día la tecnología permite disponer de esas informaciones mucho más rápido para que puedan ser utilizadas con más facilidad.

Otro tema que también mencionó María es el de la identidad profesional. Hay una confusión muy difundida entre la auditoría y la evaluación. No sé si saben de esta situación reciente, de un conjunto de científicos españoles que publicó una carta en la revista *The Lancet*, la revista médica más prestigiosa de difusión de resultados tanto médicos como de política de salud. Fue muy curioso porque la prensa, incluyendo a *El País* y otros medios, se refería a lo que habían pedido los científicos como una *auditoría externa*.¹⁵ A mí me llamó la atención porque vi que había énfasis en temas de aprendizaje, entonces fui al texto original y este claramente no se refería a auditoría sino a evaluaciones. Durante varios días los medios seguían hablando de auditorías.

Hay una confusión bastante generalizada [entre la auditoría y la evaluación], de la que se benefician los auditores pero se perjudican los evaluadores, porque los auditores en general no cuestionan los objetivos, y para los evaluadores es muy importante ver si los objetivos tienen sentido, si son relevantes, si corresponden a las necesidades del país, del sector o del Estado. OSVALDO FEINSTEIN

14 Silo, seudónimo de Mario Luis Rodríguez Cobos, escritor argentino, fundador del Movimiento Humanista.

15 Se refiere a la carta publicada el 22 de agosto de 2020, «The need for an independent evaluation of the covid-19 response in Spain» (disponible en [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)31713-X/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31713-X/fulltext)).

Hay una confusión bastante generalizada, de la que se benefician los auditores pero se perjudican los evaluadores, porque los auditores en general no cuestionan los objetivos, y para los evaluadores es muy importante ver si los objetivos tienen sentido, si son relevantes, si corresponden a las necesidades del país, del sector o del Estado. Entonces, allí hay una cuestión más profunda: no es solo ver si los recursos fueron utilizados para los fines establecidos; el trabajo del evaluador va mucho más allá. Hay una confusión de identidad y este es uno de los desafíos para los evaluadores, porque los auditores invaden el terreno continuamente.

OG —Creo que hay una pequeña confusión si creemos que los ODS son marcos de resultados. Evidentemente marcan la agenda de desarrollo para los próximos años; la gran diferencia con los Objetivos de Desarrollo del Milenio es que ponen la sostenibilidad en el centro del debate del desarrollo y era muy importante hacerlo. Hay una gran interrelación entre los Objetivos, hay superposiciones, hay múltiples relaciones, pero no son un marco coherente de resultados. Creo que hay que recoger el espíritu de la propuesta que está detrás de los ODS y ver cómo, en los múltiples niveles, se pueden hacer evaluaciones que permitan el avance en esta agenda.

[Los ODS] marcan la agenda de desarrollo para los próximos años; la gran diferencia con los Objetivos de Desarrollo del Milenio es que ponen la sostenibilidad en el centro del debate del desarrollo y era muy importante hacerlo. Oscar García

La ventaja de los ODS es que nos dan un lenguaje común para abarcar muchas áreas de desarrollo. Será trabajo de cada país identificar sus prioridades para mover la Agenda hacia adelante. Este proceso de priorización es tremendamente importante para que la evaluación pueda jugar un rol más efectivo. Estos niveles que ustedes están mencionando —local, regional, nacional, internacional, o sectoriales, o como alguna gente quiere acomodar de acuerdo a las cinco «P», para darle una mayor coherencia: *people, planet, prosperity, peace, partnership*— son formas de agregar pero, finalmente, la unidad de análisis más útil a mi modo de ver es el nivel nacional, es decir, cuáles son las políticas en un país que se dan para avanzar en la agenda de desarrollo hacia patrones más sostenibles que beneficien a su sociedad de una manera que cuente a largo plazo y que también beneficien a las futuras generaciones.

Para transitar esos distintos niveles estará siempre la pregunta que planteaba María al inicio, sobre contribución y atribución; algunos van a querer atribuirse resultados, otros simplemente van a poder contribuir a estos, y eso va a tener sus retos metodológicos. Pero creo que hay otra dificultad y es que, para poder considerar el papel de la

evaluación en fortalecer el alcance de los ODS, en este carácter de no solamente comprensión sino transformación, no bastará con un abordaje similar al que hemos tenido hasta el momento en temas de desarrollo. Necesitamos ver cuál es el tipo de acciones transformadoras que permitan dar saltos cualitativos en la aproximación a los Objetivos de Desarrollo.

Osvaldo mencionaba el importante papel de los parlamentos a nivel nacional como instrumento de *checks and balances*, de rendición de cuentas, que podrían exigir a un Ejecutivo para realmente avanzar hacia esto. Curiosamente, hasta el momento, las medidas de monitoreo y evaluación en relación con los ODS se están traduciendo en los informes voluntarios nacionales de avance hacia los ODS (*voluntary national reviews*). Estos informes voluntarios tienen muy poco contenido de evaluación; son principalmente informes de monitoreo. Creo que estos instrumentos podrían fortalecerse si comenzaran a preguntarse más el porqué y a tratar de identificar los factores que llevan a los actuales niveles de desempeño. Pero, en mi modo de ver, y esta es la propuesta que quería hacerles, estos informes podrían beneficiarse mucho más si la pregunta fuese «¿por qué no se están logrando los ODS?». Si ponemos el énfasis en por qué no se logran los ODS, va a ser mucho más fácil identificar los problemas estructurales que están bloqueando el avance hacia una agenda transformadora de un mayor desarrollo sostenible.

Déjeme ilustrarlo con un ejemplo. Uno de los Objetivos en el que trabajó muchos años el FIDA es el de cero hambre, el de erradicar el hambre. Evidentemente, la pobreza rural afecta a la mayor parte de nuestros países, los pequeños productores sufren realmente el embate de los efectos del cambio climático, de la extrema pobreza y así por delante. Si uno quiere realmente acertar en el tema del hambre, debe tener en cuenta que hay múltiples factores; no solamente la capacidad productiva de los trabajadores rurales, sino también aspectos más estructurales, entre ellos, la tenencia de la tierra y la posibilidad de comprar y vender tierra y de poder invertir en el agro de una manera rentable, etcétera. Entonces, la pregunta de por qué no se logra el hambre cero nos puede llevar a la identificación de aspectos que sí requieren cierto tipo de intervenciones de políticas públicas que puedan ser transformadoras para avanzar en la agenda.

El papel de los parlamentos de seguimiento de estos informes voluntarios de avance en el desarrollo de los ODS puede ser determinante para replicar las estructuras que una institución se daría a través los *checks and balances* internos, en que la evaluación reporta de manera independiente a un órgano de decisión como puede ser un parlamento.

JUAN PABLO MÓTTOLA (JPM) —Algo bien interesante de las tres visiones que aportaron es la cuestión de la evolución en el campo de la evaluación, de momentos iniciales en que era vista como requisito, donde quizás no había una masa crítica de evaluadores

ni de metodología ni de acumulación, a una situación actual en que tenemos ya otro camino recorrido, otra madurez en términos de formaciones, de estrategias metodológicas, de experiencia evaluativa, otra masa crítica, y que plantea un desafío hacia adelante.

Algo que también pasa en la región, en países donde todavía no tenemos procesos de institucionalización definitiva de la evaluación, especialmente en políticas públicas, es que ciertos nubarrones que se están dando a nivel económico se traducen en déficits fiscales crecientes que hacen que la toma de decisiones se vuelva muchas veces urgente en términos de presupuesto y de políticas públicas, y la tendencia

no es a utilizar la evaluación como una fuente de evidencia para tomar decisiones, sino que el sentido es otro, es el de «tenemos que reducir el déficit fiscal, el balance del Estado, llegar a cuentas fiscales equilibradas», y la evaluación pasa de ser una fuente de evidencias que ayuda a ilustrar o a ver futuros posibles, a ser vista como una fuente de restricciones al decisor. Esto es algo muy complejo, porque si efectivamente da bien, puede autoimponerle al político restricciones para tomar decisiones. Justamente, hay un contexto donde se privilegia más el balance de las cuentas públicas que otro tipo de calidad de gasto. Me surge preguntarles a ustedes: ¿qué antídoto puede haber en la evaluación para que esta situación no siga creciendo, para que ese riesgo que está arriba de la mesa se minimice? ¿Qué incentivos habría que generar o plantear para lograr que la evaluación sea vista como una fuente de evidencia virtuosa y no potencialmente como una restricción para la toma de decisiones?

OF —Creo que es interesante la experiencia del Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) en México, hasta hace unos meses —tal vez ustedes sepan que lamentablemente hubo un cambio allí—, con Gonzalo Hernández, donde se había conseguido desarrollar un sistema en el cual no solamente a nivel de



Osvaldo Feinstein

gobierno federal, sino también a nivel subnacional, había mucho interés en la evaluación. He tenido la oportunidad de participar en eventos de monitoreo y evaluación organizados por CONEVAL donde había gobernadores de varios estados de México, que tienen mucho poder, porque la mayor parte del gasto social se ejecuta con el presupuesto de los estados. Los gobernadores participaban en estos eventos y había premios, todo un sistema de incentivos, y ellos estaban muy orgullosos. Es un evento absolutamente único, en Europa no hay nada por el estilo y en América Latina tampoco. En Australia hay algo así, en menor escala. La experiencia de CONEVAL es extraordinaria y afortunadamente está bien documentada.

MB —Volviendo a lo que había planteado Juan Pablo, la experiencia española y europea en este momento con el tema covid-19 y la gestión de la crisis es que se ha empezado a generar mucho discurso sobre la necesidad de evaluar la respuesta a la crisis. Una cuestión que me parece realmente una oportunidad es que, en este momento, se reconoce —incluso lo ha hecho la Unión Europea— que la respuesta de las políticas de austeridad a la crisis del 2008 ha sido un desastre, sobre todo al agravar las desigualdades de carácter estructural. Eso está muy presente. Tiene que ver en todo, incluso cuando se está hablando ahora de los fondos de reconstrucción. No sé cómo vamos a salir, pero lo digo porque, efectivamente, en la reacción a la crisis del 2008 lo primero que se recortó fue la evaluación, en lugar de entenderla como una herramienta para recortar mejor. Desde luego, en España los recortes fueron tremendos, fueron muy lineales, muy poco basados en algún tipo de evidencia o algún tipo de estudio. Era la austeridad por la austeridad y recortar todo lo posible.

Una cuestión a la que hemos estado dándole vueltas en la evaluación de la gestión del covid-19 [es] la necesidad de evaluar con una función de reparación, es decir, dentro de la función de *accountability* de la evaluación, la idea de que hay que reparar a mucha gente en la sociedad. MARÍA BUSTELO

Me surge una cuestión a la que hemos estado dándole vueltas en la evaluación de la gestión del covid-19, y creo que es muy interesante. Lo decía la profesora Eloísa del Pino¹⁶ el otro día, en un webinar: la necesidad de evaluar con una función de reparación, es decir, dentro de la función de *accountability* de la evaluación, la idea de que hay que reparar a mucha gente en la sociedad. El covid-19 ha hecho muchísimo daño, de muchas

16 Investigadora del Instituto de Políticas y Bienes Públicos (Departamento de Economía y Política, Consejo Superior de Investigaciones Científicas, CSIC).

maneras. Me parece que es una manera interesante de enganchar con el tema de la sociedad civil. Creo que es interesante pensar en esa función, que nunca se nos había ocurrido cuando pensábamos en la *accountability*, pues estaba sobre todo la idea de la rendición de cuentas de los responsables, políticos, técnicos, etcétera. La reparación pone el foco en la sociedad, en quienes reciben las políticas públicas, no en quienes las hacen o las dejan de hacer.

El tema de los parlamentos me parece clave, pero tengo que decirles, precisamente porque he estudiado el caso español, que hay países en los que la cultura de evaluación en los parlamentos es absolutamente nula. En el caso español, todos los desarrollos que ha habido sobre la evaluación se han dado en el Poder Ejecutivo y en las administraciones públicas. Y es verdad que tenemos una función pública muy profesionalizada, pero no sé si podemos decir lo mismo respecto a los partidos políticos y los representantes en nuestro parlamento, que probablemente estén cada vez menos formados. Quiero decir que no hay ninguna cultura. Recuerdo cuando se empezó a pensar en la posibilidad de la creación de lo que luego fue la Agencia Estatal de Evaluación de Políticas Públicas y Calidad de los Servicios en España. Nos convocaron a una serie de personas expertas para pensar en cómo sería esa agencia. Por supuesto, nosotros seguíamos el libro de estilo y decíamos: «evidentemente, si esta agencia depende del Poder Legislativo, vamos a conseguir de mejor manera la independencia». No podría, porque habría que haber hecho un cambio constitucional, entonces eso incluso lo dejábamos: «que por ahora dependa del Poder Ejecutivo». Pero como se hablaba de una reforma constitucional, de una famosa ponencia del Senado que nunca llegaba a haber en el caso español, decíamos esto. Esto era en el año 2003 y, habiendo visto lo que han sido estos últimos 15 años de la actividad parlamentaria española, le hubiera visto muy poca posibilidad. Me imaginaba los informes de evaluación yendo de un lado al otro, a ver si le daban a la cabeza en el partido contrario, porque ha sido una actividad parlamentaria tremendamente histerica, terrible.

Pensando en el estudio comparado de la institucionalización de la evaluación en Europa, precisamente Suiza es uno de los países donde se puede decir que está más consolidada la función de evaluación, especialmente en el Legislativo. Curiosamente, una de las conclusiones es que hay un peligro. ¿Por qué? Porque allí se ha acabado burocratizando el tema de las evaluaciones en el funcionamiento parlamentario. Es interesante, está ahí para verlo. Yo siempre digo que es muy importante el nivel nacional, pero es muy importante también conocer los contextos. Me ha alegrado que hablaran de Chile, que efectivamente tiene mucha más cultura parlamentaria que algunos otros países.

Sobre la evaluación y las desigualdades

LF —Te quería hacer una pregunta sobre lo que planteaste respecto a que el nivel óptimo es el nivel nacional de acumulación de información sobre los ODS. ¿Qué pasa con las inequidades en un mismo territorio nacional? A países como Uruguay, Chile y otros se nos ubica actualmente como países de renta alta, sin embargo, las inequidades del territorio tiene extremos que nos llevan a trabajar de manera diferenciada, justamente para tomar en cuenta estas inequidades. ¿Cómo plantearías entonces este tema del nivel óptimo?

OG —Es un muy buen punto. Evidentemente el nivel nacional permite un nivel de rendición de cuentas mayor, porque tienes un conjunto de políticas y los datos están recolectados de manera más agregada a nivel nacional, pero esto esconde grandes desigualdades internas en todos los países. La posibilidad de afectar la agenda se da más a nivel del territorio y, evidentemente, es por eso que la participación de los gobiernos locales y de la ciudadanía en soluciones locales es fundamental. Pero vas a necesitar llegar a un nivel agregado donde puedas, al final, rendir cuentas. Por esa razón el nivel nacional es importante, si bien reconocemos que la agenda avanza mucho a nivel territorial. Pero esto, por supuesto, está sujeto a debate.

El nivel nacional permite un nivel de rendición de cuentas mayor, porque tienes un conjunto de políticas y los datos están recolectados de manera más agregada a nivel nacional, pero esto esconde grandes desigualdades internas en todos los países. OSCAR GARCÍA

OF —Quiero aportar un punto más respecto a esto. Estoy de acuerdo con la respuesta de Oscar. Me parece que esta referencia a países de ingresos altos y las desigualdades es importante. Por ejemplo, el caso de Brasil en el primer gobierno de Lula. El hecho de que Brasil no fuera un país de ingresos bajos, sino de ingresos medios-altos, permitió hacer toda una serie de políticas de distribución de ingreso y de generación de capacidades en los territorios que en un país de ingresos muy bajos no hubiera sido posible. Entonces, la evaluación a nivel nacional puede tomar en cuenta si se han aplicado políticas con intención de reducir las desigualdades y, después, cuál fue el efecto de estas políticas. O sea, se abren posibilidades que no tienen países de ingresos bajos. La distinción allí ayuda en la evaluación, al considerar las alternativas factibles.

MB —Con respecto al tema de los ODS, aquí voy a hacer mi parte desde el punto de vista del primer mundo, por decirlo de una manera nada correcta. Además de las

cosas buenísimas que aportan los ODS, además del tema de la transversalización y de la idea de que no puede haber silo en cada uno de los diecisiete Objetivos, sino que es realmente una cuestión transversal y que unos dependen de otros, que hay una interseccionalidad clara entre todos, está también el hecho de que son universales. Lo digo porque respecto a las Metas del Milenio en Europa no había ni siquiera discurso. Cuando digo universales, digo que nos hemos tenido que creer, menos mal, que también el primer mundo tenía que plantearse el cumplimiento de los Objetivos de Desarrollo Sostenible. Yo les iba a decir lo contenta que me pongo cuando veo a todos los rectores y rectoras de las universidades europeas y españolas, que ahora van todos con el pin de los ODS, y lo mismo ocurre en los ayuntamientos. Es decir que eso ha empezado ya a transmitirse.

Y me parece superimportante el comentario de Oscar sobre que habría que conseguir que los *voluntary national reviews* tuvieran más de evaluación y no sólo monitoreo. Me parece excelente la idea de por qué no se cumplen, efectivamente reconocer el carácter estructural de los problemas que tenemos, y dejar un poco de lado la ingenuidad de que en el 2030 van a estar solucionados, porque tenemos unos Objetivos muy buenos, porque, como bien decía Oscar, no están en un marco de resultados. Con esto de los *voluntary national reviews* me acordaba del Comité para Eliminación de Todas las Formas de Discriminación de la Mujer, que viene haciendo todos los años lo mismo, donde los países presentan sus informes, y una parte clarísima son los informes sombra, los informes que hace a la vez la sociedad civil, las ONG. Al final, se les ha acabado dando un hueco a esos informes sombra en la presentación de resultados, y son tremendamente importantes. De repente se me ha ocurrido que esto podría ser una manera de introducir marcos interpretativos alternativos, que no sea sólo el simple monitoreo de los países diciendo «hemos cumplido, hasta aquí hemos hecho», etcétera, y dar entrada a otros posibles actores de la sociedad civil.

LF —Para aprovechar tu *expertise* también en temas de género, me gustaría que pudieras reflexionar, en términos de balance, sobre la evaluación basada en equidad y con perspectiva de género.

MB —Creo que ha habido en la última década —no mucho más atrás— un buen desarrollo de pensar en la evaluación incorporando, integrando, la perspectiva de género. Creo que se ha avanzado mucho en ello y que todavía queda mucho por hacer. Yo siempre digo que eso realmente es el reconocimiento de que la desigualdad de género es de carácter absolutamente estructural. Es como el ejemplo que contaba Oscar sobre el hambre cero. Y eso significa que no hay ni puede haber respuestas automáticas, sino que es más bien una mirada, una forma de poner marcos interpretativos diferentes. Se ha desarrollado en los últimos años —y yo me he empeñado muchísimo en escribir sobre ello— la idea de que la evaluación puede ayudar a que esa integración de la perspectiva de género vuelva, de alguna manera, y ayude a tenérsela en cuenta en las políti-

cas públicas. Al principio, cuando se empezaba a hablar, había una idea de que, en realidad, si los programas, los proyectos, las políticas no estaban planificadas desde una perspectiva de género, iba a ser muy difícil poder evaluarlas desde una perspectiva de género. No. Al contrario. La idea es incluso no castigarnos demasiado. Yo siempre digo, en plan de risa, que los caminos del patriarcado son inescrutables. Efectivamente, es una desigualdad muy estructural y rezuma por todos lados. Entonces, lo normal o lo más esperable es que se nos rebele; es algo muy terco. Precisamente, si reconocemos eso, podemos ser capaces de darle una importancia mayor a tener esa visión y esa perspectiva en la evaluación.

Se ha avanzado muchísimo en los últimos años, especialmente porque se ha hecho mucho en políticas científicas de cómo incluir la dimensión de género en el contenido de las investigaciones, por ejemplo. Se ha avanzado muchísimo y hay mucha evidencia de los sesgos inconscientes que están operando en muchas de las políticas públicas que hacemos, como la reproducción de los roles de género, las ideas preconcebidas de género, y cómo eso opera. Entonces, la evaluación puede ayudar mucho a desvelar eso. Cada vez hay más evidencias que nos hace ver y saber dónde están los sesgos. Desde luego, no es que la pandemia haya generado desigualdades; es que las ha dejado al desnudo. No es que no existían antes, se han podido agravar, en algunos casos. Creo que es un buen momento, en el sentido de hacer caer el velo.

Se ha avanzado muchísimo y hay mucha evidencia de los sesgos inconscientes que están operando en muchas de las políticas públicas que hacemos, como la reproducción de los roles de género. MARÍA BUSTELO

El estudio que hemos hecho en la Universidad Complutense de Madrid muestra unas diferencias de género absolutamente increíbles —con diferencias estadísticamente significativas y todo absolutamente controlado— de las mujeres académicas con respecto a los hombres académicos, en todo: en impacto psicológico, en carga de trabajo, en producción académica. Creo que hay una cierta esperanza en que, cuando somos capaces de entender que hay un tema transversal y estructural que está ahí, podemos ser capaces de aplicarlo también de muchas otras maneras. No es solo el tema de género, está todo el tema de la interseccionalidad, de tener en cuenta muchas otras desigualdades y poder entenderlas, porque por sí solas no se entienden. Con esto también estoy hablando de la perspectiva medioambiental, de sostenibilidad, en todas sus vertientes.

Sobre los roles de los actores claves para el impulso de la evaluación en el marco de los ODS

JPM —Creo que es muy oportuno todo lo que se ha mencionado. Uno tiene la tentación de ver los ODM y los ODS como dos ciclos de propuestas de resultados que abarcan más allá de lo sectorial, mucho más marcado en el caso de los ODS que de los ODM; pero claramente se ve, tanto en el primer ciclo como en el segundo, que la evaluación no se está transformando en un instrumento de aprendizaje. En el caso de los ODM fue mucho más fuerte el monitoreo, pero quizás la evaluación no generó todo ese aprendizaje. En el caso de los ODS, lo que María y Oscar planteaban es muy interesante, el por qué no, o los informes sombra. Podría ser una forma de diversificar los marcos interpretativos, la metodología y demás. Junto también con lo que plantea Osvaldo sobre los roles de los distintos actores. Eso me lleva a pensar en qué roles deberían asumir los actores para lograr que el monitoreo y la evaluación se transformen en una herramienta de aprendizaje, pensando en los ODS, y no como ha estado sucediendo, que son reportes de «hasta acá llegamos» pero no pasamos a ese siguiente nivel de decir «qué aprendimos de esto, cómo podemos mejorar lo que estamos haciendo».

OF —Me parece que vale la pena destacar que lo que ha pasado con los ODM está pasando con los ODS. Es lo que mencionó Oscar sobre este tipo de informes, que tienen un potencial importante pero que, en la práctica, siguen siendo informes que simplemente van tratando de dar parcialmente cuenta de lo que está pasando. Ni siquiera llegan a ser informes completos de monitoreo, o sea, se repite el mismo tipo de problema. Entonces, es necesario ver cómo uno se desengancha de la historia, sale de esta dependencia del sendero (*path dependence*) y consigue embarcar a distintos sectores en una operación que tenga alguna posibilidad de transformar.

Ustedes listaron cuatro sectores. El término *sectores* puede generar algún equívoco pero dejemos eso de lado, por el momento. En cuanto al sector público, hay una función muy importante, en términos de recopilar y difundir información, pero también en lo que tiene que ver con la contratación de seguimiento y evaluación; no es solamente realizarlo, sino contratarlo, y contratarlo con este tipo de inquietudes, para poder no sólo rendir cuentas sino también aprender. Ni quedarse solamente con la función de aprendizaje ni quedarse sólo con la rendición de cuentas, sino combinar ambas. Y, además, todo lo que tiene que ver con la utilización de la información que va surgiendo de la implementación de las políticas.

A la sociedad civil le veo un papel fundamentalmente relacionado con la demanda de monitoreo y evaluación. También la participación en la realización, pero eso está más relacionado con lo que ustedes identificaban como otro sector, la academia. Ahí hay un

traslape, por supuesto, entre sociedad civil y academia, en todo lo que tiene que ver con métodos, con la realización de estudios y evaluaciones que puedan aportar informaciones valiosas, tanto de monitoreo como de evaluación.

El sector empresarial tiene una doble función. Una es de aportar información pero, también, en algunos casos están empezando a aportar fondos para realizar evaluaciones, sobre temas que benefician también al sector empresarial. Aquí hay toda una cuestión relacionada con los conflictos de interés, pero hay formas de hacerle frente a eso. Lo que queda fuera, y yo quería mencionarlo, es el «sector» de la cooperación internacional. No porque esté Oscar, o porque yo mismo haya estado durante tantos años vinculado a la cooperación internacional, sino porque creo que tiene tres papeles importantes. Uno es apoyar, como ha estado haciéndolo, el desarrollo de las capacidades de monitoreo y evaluación, que son muy importantes también dados los desafíos del Milenio. Un segundo papel es la evaluación de bienes públicos globales. Acá estamos hablando a nivel de país, pero hay unos aspectos de los ODS que van más allá de cada país y es importante que haya esta visión más general. Y un tercer aspecto, que me parece absolutamente central, que tiene que ver con la generación de un bien público global, el conocimiento.

En relación con la pandemia ha habido algunas afirmaciones, incluso de gente tan interesante como el historiador Yuval Noah Harari¹⁷ y Fareed Zakaria,¹⁸ quien publicó un libro recientemente sobre el tema, que dicen que lo que hay que hacer es conocer las experiencias de los países, pero no toman en cuenta la necesidad de hacer evaluaciones de estas experiencias. Es muy fácil llegar a resultados falaces cuando uno examina un país y no presta atención a aspectos del contexto, que son muy importantes. Es lo que un economista muy interesante, Paul Streeten,¹⁹ llamaba la *falacia del eclecticismo*: extraer algún elemento que parece que funcionó en un país pero que estaba asociado a otros elementos del contexto. Entonces, es importante que se realicen evaluaciones y que además se difunda el conocimiento, pero un conocimiento contextualizado. Los organismos internacionales están jugando un papel en esto, y pueden jugar un papel mucho más importante todavía en el futuro, ayudando a los países y después difundiendo esta información entre regiones y prestando la atención debida a los aspectos de contexto. Creo que vale la pena incorporar la cooperación internacional no como una quinta rueda, sino como un quinto sector, si quieren, para tener un apoyo adicional a lo que se está haciendo a nivel de país.

17 Yuval Noah Harari es un historiador y escritor israelí, profesor en la Universidad Hebrea de Jerusalén. Entre sus obras se encuentran *Sapiens: de animales a dioses*, *Homo Deus: breve historia del mañana* y *21 lecciones para el siglo XXI*.

18 Fareed Rafiq Zakaria, escritor y periodista indo-estadounidense, especializado en temas de relaciones internacionales.

19 Profesor de economía británico, nacido en Austria. Fue profesor en la Universidad de Boston (Estados Unidos). Falleció en enero de 2019.

Es muy fácil llegar a resultados falaces cuando uno examina un país y no presta atención a aspectos del contexto, que son muy importantes. Es lo que Paul Streeten llamaba la *falacia del eclecticismo*: extraer algún elemento que parece que funcionó en un país pero que estaba asociado a otros elementos del contexto. OSVALDO FEINSTEIN

OG —Quisiera profundizar un poco en el papel del sector privado. Sobre todo porque veo que ahora hay una tendencia a lo que están llamando el *impact investment*. Hay una parte del sector privado que está destinando fondos a invertir en bienes públicos ambientales como sociales que puedan tener un beneficio más allá del económico, y están utilizando métodos de evaluación, que ellos llaman de evaluación de impacto. Curiosamente, por la inmediatez de la necesidad del informe, están despreciando el abordaje tradicional de un informe de evaluación de impacto con la identificación de las líneas de base, los grupos de control, distintas técnicas, que dura un año. No les interesa. Simplemente son dedos rápidos telefónicos para ver si el impacto que se habían propuesto se alcanzó con la población beneficiaria y punto. Es una simplificación alarmante del instrumental de evaluación. Pero es una tendencia que puede crecer, así como también pueden crecer las inversiones en este ámbito de impacto social y ambiental. Creo que es importante establecer diálogos con este tipo de iniciativas, que tienen un bien común como objetivo, pero que podrían beneficiarse de la larga experiencia que existe en ámbitos académicos, de cooperación internacional y de organizaciones nacionales que vienen realizando trabajos de evaluación y es importante incluirlos en el debate.

También, como Osvaldo, quisiera mencionar un actor que no ha sido mencionado de manera muy explícita y sí fue muy bien aludido por María en sus intervenciones previas, que es el ciudadano de a pie, el ciudadano común. En mi modo de ver, ahí está la próxima frontera del ámbito de evaluación, es decir, cómo logramos la participación de las ciudadanas, los ciudadanos, que no necesariamente están organizados en un sindicato, en una asociación, en un partido político, que son nuestros primos, nuestros tíos, nuestros sobrinos, que están siendo afectados por las políticas públicas. Hoy en día el covid-19 ha acelerado el nivel de digitalización, de informatización, en nuestra sociedad; tienen acceso a canales de retorno y de *feedback plus* pero no están participando en ejercicios sistemáticos y conscientes que podrían incorporar la voz del ciudadano para fortalecer realmente este papel de la evaluación y su contribución a una democracia deliberativa.



Oscar García

Para quienes estamos en el ámbito de la evaluación, el desafío es pensar cómo podemos lograr esta participación de manera sistemática. No es una tarea fácil; es más fácil decirlo que hacerlo. Sin embargo, conjuntamente como unidad de evaluación, es importante que nos pongamos ese desafío de llegar más al ciudadano común con los resultados de evaluación, para fortalecer el papel transformador que podría tener la evaluación.

MB —Como decía antes, yo me he puesto un poco más poética, he pensado en términos más generales. Me parece que el sector público es absolutamente clave. Es donde debe desarrollarse la idea de que la evaluación es un instrumento para la profundización democrática. Es clave enfatizar la idea del desarrollo de los valores públicos y de la evaluación al servicio del interés público. En esa parte, el sector público va a tener bastante que decir.

Es cierto también el tema de la práctica de la evaluación, es decir, de encargar evaluaciones. Ahí soy más de la idea de que tenemos que pensar en sistemas de evaluación. Creo que es lo que ayuda un poco más. No solo que haya alguien perdido en algún ministerio, que tenga sensibilidad y que tenga una unidad de evaluación; no, hay que pensarlo de una manera un poco más sistémica, como sistemas públicos de evaluación, por decirlo de alguna manera.

El sector público es absolutamente clave. Es donde debe desarrollarse la idea de que la evaluación es un instrumento para la profundización democrática.

MARÍA BUSTELO

Respecto al tema de la sociedad civil, ya lo he dicho antes, creo que es muy importante y estoy con Oscar en que es realmente donde tenemos que esforzarnos muchísimo en los próximos años. Tenemos una parte ya, porque con todas las comunidades profesionales de evaluación en parte hemos llegado a una cierta sociedad civil. Pero no, es a la ciudadanía de a pie, la que decía Oscar. Para eso es muy importante que se genere un discurso de evaluación en la sociedad porque, si no, va a ser muy difícil que haya demanda de la sociedad civil.

Respecto al sector empresarial, claramente la frontera entre lo público y lo privado es más líquida y cada vez más se requiere que el sector empresarial haga actuaciones de carácter más público, que se comporte públicamente, en nombre del interés público. Sabemos que hay muchísimos problemas, como bien decía Oscar, de posibles intereses. A lo mejor, lo interesante es efectivamente llamarlos al diálogo, que entren en la conversación y no vayan por su lado.

Y con respecto a la cadena voy a hacer aquí una autocrítica importante. A mí lo que más me gusta del ámbito de evaluación es que somos una comunidad absolutamente mezclada. Por ejemplo, si yo tengo que poner en mi currículum (y me sirve mucho) que fui presidenta de la Sociedad Europea de Evaluación, en el fondo, lo que me están diciendo es «diga de qué sociedades científicas ha sido usted». Eso es lo que cuenta. Entonces, si yo les contara en qué consiste, es mucho más interesante una comunidad en la que hay gente profesional, etcétera. Desde luego, creo que es raro el académico o la académica que no venga de una reflexión práctica, que no venga de un mundo práctico anterior, como es mi caso, o que no realice trabajos profesionales a la vez en el mundo de la evaluación. Eso hace que los académicos que se dedican a la evaluación sean bastante más reales y estén más relacionados con el mundo real que los que no.

Pero es cierto que la academia tiene que seguir saliendo del ostracismo y esto también tiene que ser así en el campo de la evaluación, y creo que también en ese sentido estamos en un buen momento. Cada vez se habla más de la necesidad de que la ciencia genere un impacto social, pero todavía queda mucho. Osvaldo participó supervisando la publicación que hicieron Zenda Ofir y Thomas Schwandt, entre otros, para el Centro Canadiense para la Investigación en Desarrollo Internacional (IDRC) sobre un marco de calidad para evaluar la investigación. Cuando digo ostracismo, me refiero a ¿qué es lo que realmente está aportando la academia a la sociedad?

En ese sentido, tenemos que empezar a dejar de mirarnos exclusivamente a nosotros mismos y pensar que podemos ser evaluados también por otra gente que no sean sólo nuestros pares académicos.. En la evaluación académica, el problema es que todo el sistema está montado sobre evaluación por pares, que son más de lo mismo, es decir, no salimos. Realmente la ciencia y la academia no son evaluadas por gente que esté ahí fuera. Hay que empezar a abrir un poquito. El campo de la evaluación es ideal para hacer eso, y aquí apelo a los que son más académicos, ahí tenemos una tarea. Yo diría que efectivamente la cooperación internacional trasciende el tema del sector público en lo nacional pero, completamente de acuerdo, lo mismo tenemos que pensar en cómo los organismos internacionales, y toda la cuestión global e internacional, puede también ayudarnos a esto.

OF —Quería dejar una frase que estaba pensando, en función de lo que hemos estado hablando, que tal vez pueda tener algún sentido. «La evaluación requiere visión, observación, reflexión y pasión».

La evaluación requiere visión, observación, reflexión y pasión.

OSVALDO FEINSTEIN

LF —Totalmente de acuerdo, sobre todo lo de pasión.

JPM — Hay quien habla de evangelización, si cabe el término.

OF —Como mencioné a Silo, por ahí...

La evaluación en tiempos de covid-19: algunas lecciones aprendidas

LF —Queríamos plantear un tema adicional. Acabamos de terminar la evaluación de un plan de infancia en Paraguay con Juan Pablo y un equipo de profesionales. Como decía María, estamos en el terreno evaluando pero a la vez en la academia, y eso es mucho de nuestra profesión, a diferencia a otras profesiones y otros subgrupos de científicos. La pregunta que nos hacíamos con Juan Pablo es ¿cómo hacemos evaluaciones en contexto de covid-19?, refiriéndonos fundamentalmente al aspecto metodológico. Concretamente, este proceso de evaluación que menciono cayó en medio de la pandemia y nos planteamos los sesgos que se podrían llegar a generar en los resultados. Tuvimos

que pasar de una evaluación que iba a ser presencial, con entrevistas cara a cara y grupos focales, todo lo que ustedes ya conocen, a hacerla de manera virtual. Esta es la pregunta que queríamos plantearles.

OG —Muy pertinente y oportuna la pregunta. Déjenme compartir con ustedes algunas reflexiones. La primera es que el covid-19 ha dejado ver lo que es importante. Y todos, a nivel personal, hemos visto que lo importante es la vida, la salud de los que queremos, son las cosas esenciales: poder tener alimentación, medicamentos, etcétera. Realmente nos ha sacudido. Creo que ese sacudón también ha llegado al campo de la evaluación y por ello hemos tenido que continuar realizando evaluaciones en medio del covid-19.

El covid-19 ha dejado ver lo que es importante. Y todos, a nivel personal, hemos visto que lo importante es la vida, la salud de los que queremos, son las cosas esenciales: poder tener alimentación, medicamentos, etcétera. Realmente nos ha sacudido. Creo que ese sacudón también ha llegado al campo de la evaluación y por ello hemos tenido que continuar realizando evaluaciones en medio del covid-19. OSCAR GARCÍA

Si puedo sacar una lección de lo que hemos aprendido en este proceso de adaptación a las nuevas circunstancias, esta es el foco. Es decir, ya no podemos tratar de responder un marco de evaluación con dieciocho preguntas importantes. No hay dieciocho preguntas importantes ni tampoco tienes el tiempo ni los recursos para tratar de responder a las dieciocho. Tenemos que concentrarnos. ¿Qué es lo que quieres saber de esa determinada intervención de la niñez en el Paraguay? ¿Cuál es aquella pregunta que, si no la has respondido, realmente no hiciste el trabajo? Ese ejercicio de repensar la relevancia de la evaluación, de simplificarla y concentrar en esa pregunta es el que nos va a permitir luego pensar cuáles son los métodos para responderla. Es decir, si vamos a recoger datos a distancia, si vamos a expandir nuestras fuentes de verificación como otro tipo de *describe views*, para ver qué han hecho otras instituciones en esa misma zona, para ver si en realidad confiamos más en los datos de investigadores locales que han tratado esa problemática antes que nosotros, en fin, volvernos creativos. En mi modo de ver, entonces, la lección primera es ver qué es relevante en cada proceso de evaluación y tratar de concentrarnos ahí, porque no vamos a tener la capacidad de abarcar una evaluación como lo hacíamos antes del covid-19.

Si puedo sacar una lección de lo que hemos aprendido en este proceso de adaptación a las nuevas circunstancias, esta es el foco. Es decir, ya no podemos tratar de responder un marco de evaluación con dieciocho preguntas importantes. No hay dieciocho preguntas importantes ni tampoco tienes el tiempo ni los recursos para tratar de responder a las dieciocho. Tenemos que concentrarnos. OSCAR GARCÍA

OF — Estoy de acuerdo con lo que plantea Oscar; usaría un término que él no mencionó pero está subyacente, que es tener *selectividad* en el trabajo. Pero quisiera agregar algo. Creo que, como ha pasado en varios aspectos, el covid-19 ha profundizado y acelerado algunas cosas que ya se estaban haciendo, por ejemplo, todo lo que ya se sabe sobre la utilización de los medios virtuales, como el que estamos utilizando ahora [en este coloquio virtual]. Hay uno que me parece que tiene que ver con evaluación. Algo aludía Oscar pero quiero enfatizarlo mucho: la utilización de centros de investigación locales, nacionales. En el caso de Paraguay, por ejemplo, para ser concretos, en un trabajo como el que ustedes mencionan sería importante consultar, involucrar y hacer participar (si es que no lo han hecho ustedes ya) a una institución como el Centro Paraguayo de Estudios Sociológicos (CEPES), que está en contacto con la realidad, han hecho estudios sobre el tema y sobre temas relacionados, sobre el contexto, que pueden ayudar mucho. Porque, a la distancia, es muy fácil a veces no darse cuenta de hasta qué punto el material que se tiene está aportando realmente información que corresponde a la realidad. Creo que es muy importante, sobre todo en este momento; ya lo era antes, pero en la situación actual se profundiza todavía más e intensifica la necesidad de recurrir a expertos locales, expertos nacionales, en los temas evaluados.

MB — Me parecen excelentes las ideas de Oscar y Osvaldo sobre lo local y lo central. Desde la oficina de evaluación del PNUD están haciendo un trabajo excelente. Creo que es lo que hay; nos tenemos que adaptar, no hay otra. Yo ahora voy a hacer un estudio, un proceso metaevaluativo importante en España, que tenía toda una parte cualitativa de entrevistas y grupos focales; las entrevistas son más fácil de generar *online*, siempre y cuando se tenga la cámara, etcétera, pero en los grupos focales hay muchísimas cosas que se pierden. Pero vamos a intentarlo, luego les contaré qué tal. Aquí hay un

profesor muy bueno en temas cualitativos, Fernando Conde,²⁰ que trabajó con Jesús Ibáñez,²¹ y que se había inventado una cosa que llamaba *grupos triangulares*, o sea, grupos de a tres: no solo focalizar sino reducir la cantidad de participantes. Algunos colegas nos decían que lo habían aplicado *online* y que les permitía un poco más de control con solo tres personas. Entonces, hacer más y de forma diferente. No hay otra. La parte más cuantitativa, cuestionarios y demás, es más fácil sacarlo *online*. Y, en el tema de cómo llegar a la gente, cómo hablar con ella, cómo generar un *rapport*, vamos a tener que inventar mucho.

LF — También me refería al sesgo respecto a la información que recibimos, porque la gente a la que entrevistamos, con la que tratamos de conseguir información, te habla del pasado pero en función del contexto actual, de lo que está viviendo hoy. Ahí nos costaba mucho más tamizar desde dónde lo decían, si desde lo que realmente padecieron, si del beneficio que tuvo el plan, o no, en función de ese momento o en función de lo que están viviendo ahora concretamente.

MB — Eso habrá que incluirlo en la evaluación. Todo esto que nos está pasando es tan fuerte que no podemos pretender que no ha existido, o que no está existiendo. Entonces, eso hay que incorporarlo, claramente.

OF — Además, es bastante probable que a lo largo de estos años se hayan producido documentos que no estaban influidos por la situación actual porque fueron escritos antes, y que van dejando huellas, que se pueden tomar en cuenta y que permitirán en ese caso comparar con lo que se está planteando ahora.

OG — Estamos identificando en nuestras evaluaciones la necesidad de no solamente poner explícitamente las limitaciones, sino también lo que llamamos *reliability of the data*, o sea, cuán confiables son los datos con los que nos estamos presentando. Nos estamos haciendo esa pregunta internamente. Creo que, en la medida en que podamos ser transparentes y mencionar que estos datos son más confiables o menos confiables, de acuerdo a esto o a lo otro, porque hemos podido triangular o no, en la medida que seamos abiertos en cuanto a cómo estamos procediendo metodológicamente bajo estas nuevas circunstancias, dejamos precedentes para que otras personas que vengan luego puedan utilizar la información que estamos produciendo. Por el momento es lo que podemos hacer, ser abiertos en cuanto a las limitaciones que enfrentamos.

JPM — Agregó que, en nuestro caso, la evaluación se hacía con el aporte de evaluadores locales, nacionales, con los cuales trabajábamos asociados. Hubo dos elementos muy importantes. Uno, que era una evaluación de diseño y proceso. Tuvimos mucho cuidado con el trato y selección de personas; algunas estaban más familiarizadas con el

20 Sociólogo español. En diferentes publicaciones ha ido desarrollando una reflexión metodológica sobre la investigación cualitativa que trata de complementar esta obra desde la perspectiva del análisis del sistema de discursos.

21 Jesús Ibáñez Alonso fue un catedrático de Sociología español. Su obra teórica se centró en la metodología sociológica cualitativa y el análisis de la sociedad de consumo. Falleció en 1992.

uso de estos medios virtuales y otras no y no se podían entrevistar. Entonces hubo que apelar a toda una batería instrumental alternativa para tratar de llegar a todo el mundo y que no se nos escapara nadie, porque realmente el manejo del Zoom® y el Meet® fue una herramienta de selección importante, tuvimos que hacer mucho hincapié en eso. Segundo, cobró mucho valor (ya lo tiene de por sí) la triangulación entre entrevistas, entre métodos, y el análisis de saturación de los discursos. Eso fue muy importante para tratar de evitar que se nos metiera un sesgo sistemático dentro de nuestro análisis de proceso. Esos fueron dos elementos que nos ayudaron mucho a evitar que se afectaran de forma grave la confiabilidad y validez de los resultados.

MB —Si hay un momento en el que hay que ser más rigurosos y estrictos con la metodología, es precisamente ahora. Una de las cosas que se ha visto —unos colegas hicieron un estudio sobre esto— es la mala calidad de la mayoría de los estudios que se han hecho en estos momentos de covid-19, por ejemplo, en temas de impacto psicológico. De repente, te llegaba una encuesta para rellenar, sin ningún tipo de criterio. Yo creo que llega el momento de ser especialmente rigurosos con la metodología y reconocer las limitaciones.