

ESTUDIO TÉCNICO
N.º 07/2013

Monitoreo y Evaluación de programas
sociales en Brasil y América Latina:
conceptos, consideraciones metodológicas y
reflexiones sobre las prácticas

MDS

SAGI

Estudio Técnico

No. 07/2013

Monitoreo y Evaluación de programas sociales en Brasil y América Latina: conceptos, consideraciones metodológicas y reflexiones sobre las prácticas

Responsable por el estudio

Paulo de Martino Jannuzzi

Estudios Técnicos SAGI es una publicación de la Secretaría de Evaluación y Gestión de la Información (SAGI) creada para sistematizar notas técnicas, estudios exploratorios, productos y manuales técnicos, informes de consultoría técnica y reflexiones analíticas producidas en la Secretaría, que se ocupa de temas de interés específico para el Ministerio de Desarrollo Social y Combate al Hambre (MDS) para subvencionar, directa o indirectamente, el ciclo de diagnóstico, formulación, monitoreo y evaluación de sus políticas, programas y acciones.

El público meta de los estudios son los técnicos y responsables por las políticas y programas desarrollados a nivel federal, provincial y municipal. En esta perspectiva, son textos técnicos científicos aplicados con un alcance y dimensión adecuados a su apropiación al Ciclo de Políticas, que se caracteriza por la objetividad, el enfoque específico y la sincronización de su producción.

En el futuro, pueden convertirse en artículos para su publicación: Cuadernos de Estudios, Revista Brasileña de Monitoreo y Evaluación (RBMA) u otra revista técnica científica para llegar a un público más amplio.

Palabras claves: *Evaluación; Monitoreo; Políticas Públicas*

Unidad Responsable

Secretaría de Evaluación y Gestión de la Información

Esplanada dos Ministérios | Bloco A | Sala 307

CEP: 70.054-906 Brasília | DF

Teléfono: 61 3433-1501 | Fax: 3433-1529

www.mds.gov.br/sagi

Secretario de Evaluación y Gestión de la Información

Paulo de Martino Jannuzzi

Secretaria Adjunta

Paula Montagner

Presentación

Este Estudio Técnico presenta una discusión conceptual y metodológica sobre el campo de Monitoreo y Evaluación, basada en las prácticas en Brasil y de otros países de Iberoamérica. Trae subsidios para la estructuración de paneles de indicadores de monitoreo y encuestas de evaluación de programas sociales¹.

1. Introducción

El proceso de estructuración de Sistemas de Monitoreo y Evaluación (SM&E) de políticas, programas o, más ampliamente, de la acción gubernamental es un emprendimiento complejo, como bien y extensamente lo revela la literatura internacional sobre el tema o por cómo han aprendido – y sufrido – los gestores públicos que intentan colocarlos en marcha.

Al final, un sistema de M&E puede comprenderse, en un sentido *lato*, como el conjunto de actividades – articuladas, sistémicas y formalizadas – de producción, registro, acompañamiento y análisis crítico de las informaciones generadas en la gestión de políticas públicas, de sus programas, productos y servicios, por parte de las organizaciones, agentes y público meta involucrados, con la finalidad de subvencionar la toma de decisiones con relación a los esfuerzos necesarios para la mejora de la acción pública. Se trata, por lo tanto, de un conjunto de actividades inherentes al ciclo de gestión y de la producción de las Políticas Públicas, dirigidas a la sistematización de información con respecto a los aspectos considerados críticos para el éxito de los programas.

Además de proveer información para el análisis de las “entregas” de productos y servicios y para la corrección de eventuales fallas de la acción gubernamental, los sistemas de M&E también se destinan a generar conocimientos con respecto a los impactos de las políticas y programas, bien como de los costos de producción de los mismos. Sistemas de M&E serían, por lo tanto, dispositivos importantes para garantizar mayor transparencia de la forma de utilización de los recursos públicos y para subvencionar decisiones sobre el mérito y relevancia de las políticas y programas en la asignación presupuestaria.

Los sistemas de M&E pueden diferenciarse según varios criterios, como discuten MacDavid & Hawthorne (2006), Mackay (2007), Owen (2007), Cunill y Ospina (2008). Pueden ser más dirigidos a la mejora interna de programas, gestión y decisiones presupuestarias o rendición de cuentas de la acción gubernamental al parlamento o a la sociedad. El alcance

¹ Este texto es una revisión ampliada de otro, con el título “Estruturação de sistemas de monitoramento e especificação de pesquisas de avaliação, os problemas dos programas públicos no Brasil são”, hecho para el Curso de Evaluación de Programas en el ámbito de La Escuela Iberoamericana de Administración y Políticas Públicas, realizado en agosto de 2010 en Brasília por la Escola Nacional de Administração Pública.

conceptual y programático de los sistemas de M&E puede ser macro cuando hace referencia a políticas o estrategias de gobierno; meso cuando direccionadas al acompañamiento gerencial y analítico de programas y/u organizaciones; e inclusive microambiental cuando el énfasis de la evaluación recae sobre individuos, procesos o productos muy específicos. Pueden abarcar la integridad de acciones de gobierno, algunas áreas sectoriales o un conjunto selectivo de proyectos prioritarios.

Dependiendo de dónde esté ubicado (órgano de hacienda, órgano de planificación, órganos sectoriales, asesorías de altos dirigentes, órganos de control externo) y de la cultura existente en M&E en el gobierno, estos sistemas pueden disponer de un conjunto más amplio o restringido de herramientas y productos: mapas estratégicos, balances de gestión, sistemas informatizados, indicadores de desempeño, indicadores de políticas y programas, cuadros de indicadores, paneles situacionales, encuestas de satisfacción, diseños experimentales o casi experimentales, grupos focales de evaluación, entre otros.

Como se ve, la estructuración de sistemas de M&E involucra decisiones de orden político, administrativo y técnico. Aportar reflexiones sobre el último aspecto es lo que este texto se propone. Más específicamente, este texto busca suministrar alguna contribución de naturaleza conceptual y metodológica para ayudar en la estructuración de sistemas de monitoreo y en la especificación de investigaciones de evaluación que realmente se presten al objetivo de permitir el seguimiento continuo y el perfeccionamiento efectivo de la acción gubernamental y, más propiamente, la gestión de programas sociales, desde la perspectiva de un Ministerio sectorial.

Iniciamos con una discusión sobre los retos de legitimación y estructuración que han pasado los SM&E en Brasil y otros países. Después, hacemos una exposición del ciclo de formulación y evaluación de políticas públicas y programas, para seguidamente discutir las especificidades de los sistemas de indicadores de monitoreo y de las investigaciones de evaluación; y cómo pueden y deben integrarse.

Este texto parte de la hipótesis de que las dificultades apuntadas en el monitoreo y evaluación de programas en Brasil y probablemente en otros países de Iberoamérica —y las frustraciones derivadas— son consecuencia, en buena medida, de desarrollarlos siguiendo modelos y prescripciones muy particulares y estandarizadas, sin la debida “customización” que deberían tener en función del nivel de madurez de los programas —o del momento del “ciclo de vida” en que se encuentran—, o también de conducirlos sin un análisis previo del grado de evaluabilidad de los programas y sin una comprensión de que estas actividades se integran a procesos más amplios, e igualmente importantes, de la gestión de programas.

Forzando un poco el argumento para hacerlo más claro —aunque se perjudique el reconocimiento del esfuerzo meritorio y del trabajo competente observado en diversos sectores de la administración pública brasileña—, el hecho es que las lagunas de formación y el poco dominio de conceptos y técnicas en el campo de monitoreo y evaluación, en la

comunidad de gestores, acaban acarreado problemas en la especificación de los instrumentos de monitoreo y evaluación de las acciones gubernamentales. Esto lleva a las conocidas frustraciones con resultados tan ansiosamente esperados y a los cuestionamientos sobre la utilidad de los estudios.

Explicando mejor la motivación para la elaboración de este texto —ahora, bajo el riesgo de simplificar demasiado la naturaleza y minimizar el alcance de los estudios evaluativos conducidos por colegas investigadores en las universidades y centros de investigaciones en el país—, el enfoque marcadamente disciplinario de la encuesta aplicada en el campo, la sobrevalorización de algunos enfoques y modelos específicos de evaluación, el desconocimiento del contexto de operación de la acción pública y de la forma como los resultados de los estudios pueden ser usados más efectivamente por los gestores también contribuye para la conformación de dicho cuadro. Así, este texto busca contribuir, bajo una perspectiva modesta en substancia, pero comprometida en su sentido público, para el mejoramiento en la especificación de la demanda de instrumentos de monitoreo e investigación de evaluación por parte de la comunidad de gestores públicos y en la estructuración de los servicios que serán ofrecidos por la comunidad de investigadores académicos y profesionales.

El monitoreo y evaluación en Brasil y Iberoamérica: un campo en proceso de legitimación y estructuración

La creciente actuación del Estado brasileño y de diversos países de Iberoamérica en la promoción de políticas sociales en los últimos veinte años ha sido seguida, con algún desfase, por el perfeccionamiento de la capacidad de formulación y evaluación de programas públicos en el país. De manera general, los estudios de evaluación de demandas sociales, los diagnósticos propositivos para intervenciones específicas, los sistemas de indicadores de monitoreo y las investigaciones de evaluación de programas se han vuelto más consistentes y presentes en el cotidiano de los gestores públicos, aunque en un ritmo menos veloz que lo necesario y de forma bastante desigual por los países, por las esferas del gobierno y regiones de cada país. En el caso de Brasil, pese a los avances en la capacidad de gestión de programas, todavía son expresivos los contrastes que se pueden observar, por ejemplo, en el seguimiento de los programas en el área de salud y calificación profesional, o entre equipos de gestores del gobierno federal y técnicos de cabildos de pequeños municipios en el interior del país.

En los círculos académicos y en las distintas esferas de la administración pública hay quien dice que no es por falta de estudios, diagnósticos e indicadores que los programas sociales no consiguen alcanzar los resultados e impactos esperados. Al final, es un hecho que centros de investigación, universidades, empresas de consultoría y equipos de técnicos del sector público han producido, de forma sistemática o por medio de proyectos contratados, un amplio conjunto de trabajos, artículos y estudios sobre diferentes aspectos de la realidad social, referidos directa o indirectamente al contexto de operación de los varios programas públicos. Sin embargo, también es verdad que muchos de esos trabajos, inclusive los contratados para subvencionar la formulación de programas específicos, acaban

revelándose como diagnósticos descriptivos bastante generales, aplicables a diferentes programas sociales, con contribuciones ciertamente relevantes, pero desprovistas de informaciones más particulares y “customizadas” para formatear acciones públicas más dirigidas, en el grado que se requiere actualmente. En los diagnósticos contratados por el sector público, contribuyen para ello sin lugar a duda, las imprecisiones o lagunas de los términos de referencia —que reflejan muchas veces la falta de claridad de los objetivos y la naturaleza del programa público a ser implementado, hecho derivado, a su vez, de las deficiencias de formación del gestor público— y la escasez de tiempo y recursos para la realización de levantamientos de campo específicos. En este contexto, acaban justificándose estudios de diagnósticos más generales, basados en trabajos anteriores y en fuentes de datos e investigaciones ya disponibles.

También es necesario reconocer las limitaciones de las investigaciones de evaluación de programas públicos como fuentes de información para la readecuación de los programas existentes o para la propuesta de nuevos. En las evaluaciones de impacto, las constataciones empíricas acerca de la efectividad de los programas son usualmente ambiguas o poco consistentes. En los casos en que se adoptan delineaciones metodológicas casi experimentales, las condiciones para garantizar la validez interna de los estudios acaban por inviabilizar la apropiación y generalización de los resultados para evaluaciones más amplias de programas (WORTHERN *et ál.*, 2004).

Varios estudios evaluativos enfocados en el proceso de implementación de los programas acaban teniendo bajo poder de generalización debido a los vacíos de conocimiento sobre las actividades desarrolladas y el papel de los agentes institucionales involucrados en los programas, sea por las limitaciones o inexistencia de sistemas de información de monitoreo, sea por las decisiones metodológicas en cuanto a las técnicas, muestras y casos investigados, muchas veces escogidos más en función de los plazos y recursos disponibles que por las cuestiones a las cuales se quiere responder. Evaluaciones dirigidas al análisis de la consistencia teórica y del diseño lógico de programas —que podrían producir conocimiento más generalizable en el área— son menos frecuentes o, por lo menos, menos conocidas que otros tipos de estudios evaluativos en el país.

En fin, pese a los avances en el área, es necesario reconocer que la insatisfacción con los estudios evaluativos todavía es más frecuente que lo deseable. Como ya lo revelaba Cotta hace más de diez años sobre el caso brasileño:

Las metodologías de evaluación de programas sociales han sido objeto de severas críticas. Básicamente, se afirma que en la práctica las evaluaciones no subvencionan el proceso decisorio porque sus resultados son inconcluyentes, inoportunos e irrelevantes. Inconcluyentes en función de las propias limitaciones de este tipo de estudio, inoportunos debido a la morosidad del proceso evaluativo, e irrelevantes porque no responden a las demandas informativas de todos los agentes sociales sujetos a la intervención. (COTTA 1998, p.118).

Una situación semejante parece haber atravesado los Estados Unidos (EEUU) en los años 1970, a juzgar por el reporte del General Accounting Office que, movilizadado por demanda del Senado americano en 1974, constató que las informaciones de las investigaciones sociales y estudios evaluativos financiados con recursos públicos no contribuían para el diseño de las políticas públicas, por la dispersión y fragmentación de temas investigados, falta de coordinación y orientación de lo que era de interés gubernamental investigar, dificultad de acceso y comprensión de los reportes de investigaciones. (GAO, 1977).

Aun así, recientemente se han constatado limitaciones y problemas en los estudios evaluativos de programas públicos en EEUU, a veces con despliegues hasta peores que los descritos por Cotta. De hecho, como coloca Worthern *et ál.* (2004):

De vez en cuando, una "evaluación" mal concebida o mal ejecutada produce informaciones que, en el mejor de los casos, serían engañosas y, en el peor, absolutamente falsas. Aunque esto sea poco común, pueden causar problemas graves. Como generalmente tienen aire de respetabilidad, dichas evaluaciones no acostumbran ser cuestionadas, y el resultado es que decisiones importantes sobre programas y servicios esenciales se basan inadvertidamente en informaciones falaces. (WORTHERN *et ál.*, 2004, p.44).

Conspiran para la conformación de dichos problemas, entre tantos factores, la creencia desmesurada en la capacidad de anticipación e implementación de programas por parte de técnicos de alto rango que al no incorporar la contribución de agentes involucrados en el trabajo, acaban por diseñar procesos y rutinas que desconsideran las distintas realidades de operación de los programas. Este tecnocratismo ingenuo invariablemente se olvida de las dificultades de articulación y colaboración de los tres niveles de gobierno o de carteras sociales en una misma esfera y de las diferencias de capacidad de gestión y control social fuera del país. Tan o más grave, es que también desconsidera la heterogeneidad socioeconómica de la población demandante en la formulación de los programas, considerándola como única población meta, para lo cual se debe proveer un mismo conjunto de servicios, de norte a sur, del centro a la periferia de las ciudades. Así, con problemas en las fases iniciales del ciclo de gestión de programas —en el reconocimiento de las cuestiones sociales y en el diseño de las intervenciones idealizadas para mitigarlas— no se podrían esperar, de hecho, resultados impactantes.

Desconocimiento sobre el nivel de evaluabilidad de los programas es otro factor que mina la credibilidad de los instrumentos de monitoreo y evaluación. La prematuridad en la encomienda de estudios evaluativos de resultados e impactos, cuando se sabe que el programa todavía se encuentra en fase de implantación o con problemas de gestión; y la anticipación de evaluaciones externas, con naturaleza más de auditoría en detrimento de evaluaciones de carácter más formativo, conducidas internamente y dirigidas al perfeccionamiento incremental del programa son problemas más frecuentes de lo que se podría esperar. Son aportados volúmenes significativos de recursos en investigaciones de evaluación, procurando garantizar representatividad muestral de resultados a nivel

nacional, olvidándose de estructurar paneles o sistemas de indicadores de monitoreo que permitirían identificar —con todas las limitaciones de cobertura, pero con la tempestividad debida— buena parte de los problemas a tiempo de ser corregidos.

Efectivamente, uno de los hallazgos sistemáticos de las evaluaciones de programas realizadas por el Tribunal de Contas da União² en Brasil hasta hace algunos años es la inexistencia de sistemas de información para seguimiento de las actividades de los programas. En este sentido, es ilustrativa la constatación, reproducida a seguir, acerca de la evaluación del Programa Nacional Biblioteca en la Escuela (PNBE) operado por el Fondo Nacional de Desarrollo Escolar (FNDE) del Ministerio de Educación (MEC), con el objetivo de distribuir un acervo de libros para formación de bibliotecas en las escuelas.

El trabajo constató que el FNDE ha mostrado una gran eficacia operativa en la distribución de los acervos del PNBE para las escuelas beneficiadas. Fueron atendidas 20 mil escuelas en 1998 y 36 mil en 1999, con perspectiva de atender 139 mil escuelas en 2002.

Además, también se verificó que el PNBE no tiene actividades de monitoreo y evaluación bien estructuradas. El MEC no dispone de informaciones que permitan conocer el nivel de utilización de los acervos, así como los problemas que pueden estar afectando la efectividad del uso de los libros en el desarrollo escolar de los alumnos en las escuelas beneficiadas. Este hecho es más preocupante justamente en el momento en que aumenta la escala de atención del PNBE (TCU, 2002, p.10).

Para ello, contribuye el nivel aún incipiente de conocimiento del campo aplicado de monitoreo y evaluación de programas en Brasil, sea dentro de la estructura pública, sea en las universidades y en los centros de investigación. Mientras que en EEUU y en países europeos desarrollados la cultura de evaluación de programas ya se encuentra en una etapa maduraⁱ, después de más de tres o cuatro décadas de desarrollo, con marcos conceptuales amplios y probados, con profesionales con formación multidisciplinaria y con instituciones especializadas, en Brasil y otros países de Iberoamérica la institucionalización del campo es mucho más recienteⁱⁱ.

Como observan Worthern *et ál.* (2004), la cultura de monitoreo y evaluación de programas en los Estados Unidos —así como los instrumentos de planificación y programación presupuestaria— comenzó a fortalecerse en los años 1960 en el gobierno Lyndon Johnson, cuando se estructuraron varios programas sociales a nivel federal, en el contexto de lo que se denominó Great Society, y de la necesidad, por lo tanto, de garantizar una mejor gestión de los programas públicos. Rossi *et ál.* (2004) identifican esfuerzos sistemáticos en la evaluación de programas sociales norteamericanos, a partir de la década de 1930, como despliegue del papel creciente del Estado americano en la financiación de servicios sociales (respondiendo a los efectos de la crisis de 1929). La evaluación de programas, que se concentraba inicialmente en las áreas de educación

² Tribunal de Contas es la institución de Control Externo subordinado al Parlamento brasileño.

(sobre todo en la investigación de programas de alfabetización), de salud pública (en las acciones de combate a la mortalidad por enfermedades infecciosas y parasitarias) y de calificación profesional, en la posguerra pasa a abarcar áreas como prevención de la violencia juvenil, planificación familiar, nutrición, programas habitacionales, desarrollo rural.

Sin embargo, en los años 1960 el campo de la evaluación de programas pasa por un "boom", con la ampliación de los programas sociales, como ya fue mencionado y como indican Rossi *et ál.* (2004), con el avance y diseminación de las técnicas aplicadas de la investigación social, desarrolladas en las universidades americanas.

Al contrario de lo que se podría imaginar, en las décadas siguientes, el énfasis en la desreglamentación y descentralización de la prestación de servicios públicos y en el cuestionamiento sobre la pertinencia y tamaño de los programas sociales en los gobiernos republicanos conservadores, que asumieron el poder más tarde, acabó por valorizar la práctica de la evaluación, al diseminar la cultura del monitoreo y la evaluación para gobiernos provinciales, municipales y organizaciones no gubernamentales. Al final, inclusive un "estado regulador" requiere instrumentos de medición de resultados para dar seguimiento a la prestación de los servicios financiados con recursos públicos, inclusive para justificar, con alguna base técnica y transparencia, la discontinuidad de programas sociales "ineficaces" y redirigir los gastos para otras áreas (como hacia los gastos militares de la guerra de Vietnam, en el gobierno Nixon, y para defensa militar en el gobierno Reagan). Respondiendo a estas demandas, comentan los autores que en los años 1970 fueron estructurados cursos de postgrado de naturaleza disciplinaria y más tarde multidisciplinaria con fuerte énfasis en técnicas de investigación (cuantitativas y cualitativas), abriendo la posibilidad de crear una comunidad profesional de evaluadores¹.

La evaluación de programas dejaría de ser preocupación apenas para científicos sociales, en temas de investigaciones académicas en las universidades, ganando proyección e interés en la comunidad de gestores públicos como instrumento para el perfeccionamiento de los servicios públicos. Este cambio cualitativo de la producción y apropiación de conocimientos en el área —que parece procesarse en este momento en Brasil— es muy bien registrado por la siguiente referencia de Rossi *et ál.* (2004):

In its early years, evaluation was shaped mainly by the interests of social researchers. In later stages, however, the consumers of evaluation research exercised a significant influence on the field. Evaluation is now sustained primarily by funding from policymakers, program planners, and administrators who use the findings and by the interests of general public and the clients of the programs evaluated. Evaluation results may not make front-page headlines, but they are often matters of intense concern to informed citizens, program sponsors, and decisionmakers, and those whose lives are affected, directly or indirectly, by the programs at issue (ROSSI *et ál.*, 2004, p. 9).

En el caso brasileño, actualmente, una parte significativa de la evaluación de programas públicos ha sido realizada por equipos de centros de investigación y universidades, con buena experiencia en análisis de macropolíticas, coyuntura social o proyectos de investigación académica, pero en general sin experiencia en la implementación real de programas y sin reunir el conocimiento multidisciplinario requerido para abordar los problemas complejos donde los programas buscan actuar². No hay dudas que, en general, estas evaluaciones de programas cuentan con equipos competentes en el empleo de las buenas prácticas de investigación social y en el uso riguroso de técnicas cuantitativas y cualitativas de recolección y análisis de datos.

Como en las investigaciones académicas, es fundamental garantizar el rigor metodológico en los estudios evaluativos, para permitir un conocimiento más circunstanciado de los problemas, las soluciones idealizadas para mitigarlos y para conferir legitimidad técnica-científica en la evaluación de la acción gubernamental. Pero, como ya había sido alertado en uno de los primeros manuales de gran amplitud de evaluación de programas a inicio de los años 1970, programas sociales son "objetos de investigación" inherentemente difíciles de ser tratados, por el ambiente muchas veces inhóspito y poco colaborativo, por el contexto político menos o más favorable, por la naturaleza evasiva de las preguntas y cuestiones que son demandadas para que el evaluador responda, por la menor o mayor disponibilidad de datos, por el tiempo y recursos disponibles para conducir análisis más consistentes (Weiss, 1972). En manuales más recientes, como los citados de Worthern *et ál.* (2004) y Rossi *et. ál.* (2004), se recomienda que los estudios evaluativos sean realizados por investigadores con conocimiento técnico y metodológico apropiado con el tema en cuestión, pero también con experiencia empírica anterior y, sobre todo, con desprendimiento para innovar e improvisar en las diversas circunstancias y dificultades que insisten en aparecer en el cotidiano práctico del área. Así, con frecuencia, estudios evaluativos realizados por "puristas metodológicos" llegan a resultados ya conocidos por el gestor de programas, o concluyen — con ingenua asertividad— acerca de la necesidad de discontinuar el programa evaluado, por la supuesta baja efectividad establecida, según sus selecciones metodológicas sobre qué investigar y cómo hacerlo.

Estas consideraciones iniciales ilustran las dificultades o deficiencias de los sistemas de monitoreo y los frecuentes problemas de especificación de las investigaciones de evaluación para dar seguimiento a las acciones y programas gubernamentales en Brasil y posiblemente en otros países de Iberoamérica. Muchos son los ejemplos de investigaciones de evaluación que, aunque son motivadas por preocupación legítima y meritoria con la eficacia y efectividad del gasto público, revelan infelizmente resultados triviales, metodológicamente cuestionables o poco apropiados para la reformulación de los programas. La búsqueda de soluciones es, más allá de la toma de decisiones políticas y administrativas, profundizar el conocimiento sobre los aspectos técnicos de los sistema de M&E. ¡Vamos a ellos!

El ciclo de políticas públicas y programas

Antes de pasar a una explicación sobre el ciclo y sus etapas, vale formalizar dos términos citados normalmente en el texto: políticas públicas y programas. Esto es importante para circunstanciar los límites del campo de diálogo establecido, ya que evaluación de políticas públicas y evaluación de programas públicos son términos muy superpuestos, pero se refieren a contextos muy diferentes (en amplitud) del análisis de la intervención pública.

Evitando entrar en una discusión muy extensa sobre los diferentes significados del término —sistematizados en Villanueva (2006)— política pública es, según la definición de Nascimento (1991), el conjunto de decisiones tomadas por las instituciones del Estado buscando la solución de un problema o para la reconducción de una tendencia, con la intención de orientar su evolución para un fin establecido como deseable. Como aclara Saravia (2006), en

[...] una perspectiva más operativa, podríamos decir que la política pública es un sistema de decisiones públicas que procura acciones u omisiones, preventivas o correctivas, destinadas a mantener o modificar la realidad de uno o varios sectores de la vida social, por medio de la definición de objetivos y estrategias de actuación y de la asignación de los recursos necesarios para alcanzar los objetivos establecidos (SARAIVA, 2006, p.29).

Las decisiones que conforman dicha política pública —distributiva, redistributiva o regulatoria; universal o focalizada— asumen diferentes formatos: pueden ser enunciados de directrices estratégicas de gobierno, leyes y decretos normativos, que especifican de forma más clara y operativa dichas directrices, creación de organizaciones o programas que van a conceder operatividad a las directrices, leyes y normas (VILLANUEVA, 2006).

El programa público es uno de los instrumentos operativos de las políticas públicas. Más específicamente, se trata de un conjunto sistémico de acciones programadas y articuladas entre sí, con el objetivo de atender a una demanda pública específica, asentada en la agenda de prioridades de políticas públicas del Estado o Gobierno. Así, un programa social es un conjunto de actividades dirigidas a solucionar un problema vivido por la sociedad, en su todo o por grupos (ROSSI *et ál.*, 2004).

Evaluación de políticas públicas, o mejor, análisis de políticas públicas, término preferido por Owen (2007), es una actividad mucho más amplia que la evaluación de programas. Trata del contexto político-social del surgimiento de la política, de los actores participantes. Se dirige más hacia la aclaración de su proceso de construcción que hacia la recomendación práctica de cómo perfeccionarla, constituyendo en realidad un campo de investigación más propiamente académico que la perspectiva técnica-profesional en que se realiza la evaluación de programas. Se debe observar que hay programas, inclusive, con objetivo temático y escala de operación mucho más amplia que lo usual, revelándose más como un paraguas de otros programas más específicos. Otros programas pueden ser proyectos de

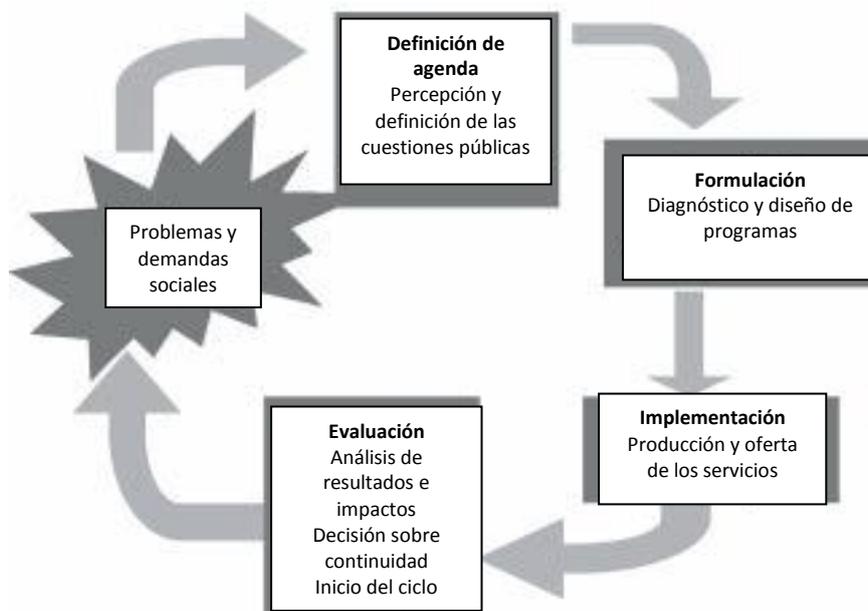
corto alcance, o actividades que tienen como objetivo la elaboración de un producto o servicio finalizado o en proceso. La discusión aquí instaurada es suficientemente amplia para contemplar estas modalidades.

En los manuales clásicos de Ciencias Políticas, el proceso de formulación de políticas públicas ha sido presentado periódicamente por el ciclo de etapas sucesivas (*Policy Cycle*), con más o menos niveles, como ilustrado en el Diagrama 1. Donde, pese a las críticas de mucho tiempo, cuanto a la forma simplificada con que este diagrama presenta el proceso político y su propia veracidad empírica, la separación en etapas se presta a los objetivos de evidenciar, durante el proceso, énfasis diferenciados en la planificación, operación o evaluación de los programas. Además, se justifica para fines didácticos y para orientar el recorte analítico en la investigación académica en el área (NASCIMENTO, 1991) ⁱⁱⁱ.

En este modelo, la primera etapa —Definición de la agenda política (*Agenda-setting*) — corresponde a los múltiples caminos y procesos que culminan con el reconocimiento de una cuestión social como problema público y de la necesidad de la acción gubernamental para su solución; es decir, la legitimación de la cuestión social en la pauta pública o agenda de las políticas públicas del país, en determinado momento. La siguiente etapa —Formulación de políticas y programas (*Policy formulation*) — se refiere a los procesos y actividades relacionadas a la construcción de posibles soluciones, orientaciones y programas para lidiar con el tema recién legitimado en la agenda. Entonces, siguiendo la secuencia, es necesario seleccionar —en la Toma de decisiones técnica-política (*Decision making*) — , el rumbo a ser seguido, de acción efectiva o no, decidiéndose por una o algunas de las alternativas formuladas. La cuarta etapa —Implementación de políticas y programas (*Policy implementation*) — corresponde a los esfuerzos de implementación de la acción gubernamental, en la asignación de recursos y desarrollo de los procesos previstos en las alternativas y programas escogidos anteriormente. Por último, es necesario analizar si los esfuerzos emprendidos están actuando en el sentido esperado de solucionar el problema original —etapa de Evaluación de las políticas y programas (*Policy evaluation*). Es necesario evaluar si es necesario realizar cambios en los programas implementados para garantizar su efectividad; discontinuarlos, si el problema dejó de componer la agenda; o entonces adaptarlos a una nueva realidad, reiniciando el ciclo.

Es oportuno registrar que la evaluación, como **etapa** del ciclo, se realiza después de la implementación. Se trata de un momento de naturaleza más reflexiva para la continuidad o no del programa. Por lo tanto, se distingue de las **actividades** de monitoreo y evaluación que se realizan mediante los sistemas de indicadores y las investigaciones de evaluación, instrumentos investigativos que pueden ser empleados en cualquier momento del ciclo, como será discutido más adelante. Denominar esta etapa decisiva del ciclo como de **evaluación sumativa** probablemente ayude a evitar el doble sentido que el término asume en el área.

Diagrama 1: El ciclo de políticas y programas públicos



La agenda política corresponde al conjunto de asuntos y problemas que los gestores públicos y la comunidad política entienden como siendo más relevantes en dado momento y, no necesariamente a la lista de preocupaciones de la sociedad o titulares de la prensa (JANN; WEGRICH, 2007). Al final, el reconocimiento de una cuestión social como problema de Gobierno o Estado no es un proceso simple e inmediato, que responde automáticamente a las estadísticas disponibles, por más reveladoras que sean de la gravedad del tema, cuando es comparada a otros países o a otros momentos del pasado. No es la voluntad de un técnico del sector público, un investigador académico o gobernante electo, con conocimiento empírico consistente de la realidad o visión osada, que garantiza inmediatamente su incorporación en la agenda formal de gobierno. Como coloca Parada (2006, p.73), "no toda idea entra a la agenda. No todos los temas de la agenda se convierten en programas".

Cuando existe la convicción de que un problema social necesita ser dominado política y administrativamente, es cuando se transforma en una cuestión pública (FREY, 1997). Si parece haber una solución técnica viable y factible para determinada cuestión social, ésta entra más fácilmente en la agenda. Al final, la estructura del sector público, por los mecanismos institucionales existentes y operantes, es un ambiente que "digiere" innovaciones a su propio tiempo y estilo¹.

Existen varias interpretaciones diferentes de cómo es construida la agenda. En todas las perspectivas es entendida como un proceso colectivo y conflictivo de definición, como señalan Howlett & Ramesh (2003). Va siendo formada por medio de diferentes mecanismos de presión externa de grupos organizados en partidos políticos, sindicatos, asociaciones patronales,

prensa y otras instituciones, por los temas por ellos entendidos como relevantes; por iniciativa interna del gobierno y movilización social en función de compromisos asumidos en las elecciones; por la actuación de la burocracia pública en la conducción de las demandas sectoriales, etc.

Los actores políticos son elementos catalizadores de este proceso. Son diversos y poseen características distintas: políticos y burócratas; empresarios, trabajadores / sindicatos, servidores públicos y los medios de comunicación (RUA, 1998). Grandes empresarios, individualmente o por medio de sus *lobbies* (grupos de presión), remiten sus demandas y presionan a los actores públicos, en las decisiones gubernamentales, a favor de sus intereses. Otro actor político de gran importancia son los agentes internacionales, como el Fondo Monetario Internacional (FMI), Banco Mundial, Organización de las Naciones Unidas (ONU), Organización Mundial del Comercio (OMC), etc.; o los países con los cuales Brasil mantiene relaciones políticas. Estos actores pueden afectar no apenas la economía, como también la política interna del país (RUA, 1998). Además, los medios de comunicación son otros agentes importantes en el proceso — especialmente la televisión, con difusión masificada en todos los estratos de la población— puesto que disponen de recursos para influir la opinión pública en la formación de las demandas (FREY, 1997). En fin, las organizaciones políticas —partidos, sindicatos, grupos de interés— son fundamentales para que las demandas entren en la agenda política del gobierno y así, puedan transformarse en acciones y programas concretos.

La amplitud de la agenda política estaría condicionada también al nivel de desarrollo económico y tecnológico del país, puesto que dichos factores, al viabilizar aumentos crecientes de productividad en la economía, permitirían atender a las demandas de grupos de interés organizados, por la ampliación de la oferta, cobertura y diversidad de los servicios sociales (CAREY, 2006). El ritmo de constitución de dicha agenda de bienestar social sería potencializado, en contextos de mayor libertad y participación política, con mayor actuación de partidos políticos, sindicatos, prensa y otros grupos de interés. Los niveles de urbanización y envejecimiento poblacional también serían otros factores condicionantes importantes para entender la formación de la agenda política, por la visibilidad y concretización que confieren a los problemas sociales derivados.

En la etapa de la formulación de la política, los problemas, propuestos y demandas explicitadas en la agenda se transforman en leyes, programas y propuestas de acciones. Es decir, la formulación de política involucra la búsqueda de posibles soluciones para los temas priorizados en la agenda. En este momento, los elementos operativos de la política en cuestión necesitan ser explicitados: directrices estratégicas, propuestas de leyes, decretos normativos, programas y proyectos. En una visión simplificada —y un tanto romántica e ingenua, como diría Charles Lindblom— del proceso, dichos programas y propuestas serían elaborados por el empleo de técnicas generales de planificación de proyectos; conocimientos sectoriales específicos; análisis de viabilidad política, de costo-beneficio o costo-efectividad de las soluciones; revisión crítica de experiencias anteriores y buenas prácticas identificadas. Estas tareas son conducidas por técnicos del sector público, con mayor o menor participación de consultores externos,

siguiendo orientaciones emanadas de gestores públicos en el tope de la jerarquía.

En función de la naturaleza del problema y orientación general de gobierno, los programas propuestos se clasificarían en acciones de cuño distributivo, redistributivo o regulatorio (HOWLETT y RAMESH, 2003). En realidad, esta fase no está tan desvinculada del cotidiano de la política o tan aislada de las influencias de las propuestas de los grupos de interés, como se supone. Muchas veces, las propuestas de soluciones llegan junto con las presiones para introducir el tema en la agenda formal de gobierno. Si no es colocado el tema, siempre habrá actores con mayor o menor influencia en la formulación de los posibles programas y acciones (JANN; WEGRICH, 2007).

Establecidas las diferentes alternativas para un tema social, es necesario seleccionar cuál deberá ser adoptada. Mientras que en la definición de la agenda política y formulación de programas la participación de agentes no pertenecientes al gobierno o Estado desempeña un papel central, la toma de decisión involucra a los gestores y técnicos más directamente relacionados a la política o al programa. Este es el costo y beneficio del ejercicio de la actividad pública. Esto no significa que otros agentes y actores no estén activos en el proceso, de manera que influyan en la decisión en uno u otro sentido, introduciendo —legítimamente en muchos casos— consideraciones de orden política en rutinas de carácter más eminentemente técnico. Como bien coloca Howlett & Ramesh (2003):

These actors can and do, of course, engage in various kinds of lobbying activities aimed at persuading, encouraging, and sometimes even coercing authoritative office-holders to adopt options of which they approve.

However, unlike office-holders, those other actors have, at best, a voice in the decision-making process, but they do not have a vote per se (HOWLETT y RAMESH, 2003, p.163).

Entonces, llega el momento de la implementación de los programas, proyectos y acciones puntuales o más amplias. La fase de implementación es el momento en que la política pública comienza a ser ejecutada, en que los recursos asignados comienzan a transformarse en acciones capaces de actuar y cambiar cierta realidad. Aquí la política adquiere una mayor concretización objetiva y, por consiguiente, comienza a revelar más explícitamente las dificultades de su introducción y operación por los agentes encargados de “meter la mano en la masa” (NASCIMENTO, 1991). Se trata de la gestión de las acciones para hacer que una política salga del papel y funcione efectivamente, comprendiendo el conjunto de acciones realizadas, por grupos o actores de naturaleza pública o privada, para la consecución de objetivos definidos. Involucra programación de actividades, previsión de recursos financieros, asignación de recursos humanos, movilización de agentes, interlocución con actores estratégicos, manejo de los mecanismos que aseguren la gobernabilidad de las actividades y corrección de trayectorias, en caso que se identifiquen obstáculos o sorpresas no anticipadas.

Es en la evaluación —o mejor dicho, en la evaluación sumativa— que son confrontados los resultados esperados con los resultados alcanzados. Es el

momento en que se da un retorno a la etapa inicial de formulación, con correcciones en el plan de acción, en caso que sea una política de orden continuo; o cuando se decide por la conclusión de la política. Es también una etapa de verificación de los impactos deseados e indeseados. Se trata de verificar los efectos obtenidos para deducir consecuencias indeseadas en las acciones y programas futuros. Como bien coloca Frey (1997), la evaluación es imprescindible para el desarrollo y la adaptación continua de las formas y de los instrumentos de acción pública, lo que puede ser denominado como la fase de "aprendizaje de la política".

El ciclo de políticas públicas tal como fue descrito es, ciertamente, un modelo idealizado, con adherencia bastante limitada al cotidiano verificado en Brasil o en otros países. Como bien observa Saravia (2006, p.29), "el proceso de política pública no posee una racionalidad manifiesta. No es una ordenación tranquila en la cual cada actor social conoce y desempeña el papel esperado". Por lo que sugiere Lindblom (2006), la actualización de la agenda, formulación de políticas y programas, implementación y evaluación serían etapas permanentes y concomitantes del proceso político. La actualización de la agenda política y las innovaciones de los programas serían, de hecho, mucho más incrementales y continuas que lo que les gustaría a los gestores públicos. Tampoco habría, según Parada (2006, p.72), un momento de evaluación cabal de las políticas y programas, siendo "[...] más habitual que cambien o se combinen con otras. Se ha llegado a decir que las políticas públicas son inmortales".

Buena parte de la crítica hecha a este modelo se refiere al hecho de sugerir que la administración pública, sus gestores, los actores políticos y los técnicos actúen de forma bastante sistemática y cooperativa, como si estuviesen todos involucrados en la resolución de un problema consensualmente percibido, empleando métodos racionales y objetivos en la búsqueda de la solución, de acuerdo con una secuencia lineal de etapas bien trazadas. Como diría Lindblom (2006), una vez más, el imperativo de la racionalidad técnica en el proceso, el análisis exhaustivo de los problemas, la búsqueda de soluciones óptimas y la creencia en el poder revelador y conciliador del discurso técnico-científico —tan presente en los entes de planificación pública— no serían empíricamente constatables, ni factibles y mucho menos deseables.

En realidad, la formulación de políticas se configura como un proceso que involucra la interacción de muchos agentes, con diferentes intereses. Es marcada por apoyos entusiasmados de algunos, resistencias legítimas o no de otros. Está repleta de avances y retrocesos, con despliegues no necesariamente secuenciales y no plenamente anticipables. Como observan Jann & Wegrich (2007):

Policy process rarely features clear-cut beginnings and endings. At the same time, policies have always been constantly reviewed, controlled, modified and sometimes even terminated; policies are perpetually reformulated, implemented, evaluated and adapted.... Moreover, policies do not develop in a vacuum, but are adopted in a crowded policy space that leaves little space for innovation [...] (JANN; WEGRICH, 2007, p.44-45).

No obstante, en la visión de Howlett & Ramesh (2003), este modelo de representación tiene la gran virtud de facilitar el entendimiento del complejo proceso de interacción de diversos agentes, en las múltiples etapas por las que pasa la formulación de políticas públicas, ofreciendo un marco metodológico general para análisis aislados de cada etapa del proceso o de las relaciones de cada una con las demás, hacia adelante o hacia atrás. Este modelo sería también suficientemente general para ser aplicado en el entendimiento del proceso en la mayoría de los ámbitos y contextos de formulación de políticas y programas —a nivel federal, provincial, local e inclusive sectorial.

Jann & Wegrich (2007) agregan que el éxito y la “elasticidad” de este modelo clásico a las críticas de investigadores académicos es que se parece a un modelo prescriptivo ideal a ser adoptado, donde gestores electos gobiernan con cuerpos burocráticos con buena capacidad de diagnosticar problemas, con posiciones activas e innovadoras en la formulación de programas, operando con racionalidad técnica en la toma de decisiones, con control efectivo de las actividades en la implementación de los programas públicos y objetividad en su evaluación.

Los sistemas de indicadores de monitoreo en el ciclo de políticas y programas

La preocupación con la construcción de indicadores de monitoreo de la acción gubernamental es tan o más antigua que la propia evaluación de programas públicos, si fuesen consideradas las medidas de *performance* más general del sector público basadas en la entrega de obras y productos y en el cómputo de indicadores de costos de servicios, a partir de la ejecución presupuestaria. Como registran McDavid y Hawthorn (2006) en el libro que busca ofrecer una visión integrada entre evaluación de programas y medición del desempeño gubernamental:

While we have tended to situate the beginnings of performance measures in the United States, in the 1960s, with the development of performance management systems such as programmed planned budgeting systems (PPBS)[...] there is good evidence that performance measurement and reporting was well-developed in some American local governments early in the 20th century (MCDAVID; HAWTHORN, 2006, p.283).

Los autores citan las iniciativas de algunos gobiernos locales en los EEUU, antes de la Segunda Guerra Mundial, en la elaboración de reportes con cómputo regular de medidas de eficiencia de los servicios públicos entonces ofrecidos, como la construcción y mantenimiento de las calles y carreteras, distribución de agua y recolección de basura. Los productos y servicios puestos a disposición por el sector público municipal eran bastante simples y tangibles, así como la contabilización de los costos para su producción.

Pero es la introducción de las técnicas de programación presupuestaria por programas en el gobierno Kennedy que marcaría un momento de mayor perfeccionamiento en la medición del desempeño —expresado por indicadores de eficiencia y efectividad— en la esfera federal americana. Al proponer la estructuración del presupuesto según el conjunto de actividades

y programas afines —ahora mucho más complejos que los ofrecidos por los municipios en la primera mitad del siglo XX— y no según las agencias y organizaciones que las ejecutaban, se esperaba conseguir definir los resultados y costos más consistentes de los servicios públicos disponibles y, por lo tanto, medidas de eficiencia más precisas (MCDAVID; HAWTHORN, 2006).

Pese al fracaso en la implementación del presupuesto-programa, en el vínculo entre recursos-productos a nivel de programas, y el contexto de crisis fiscal del Estado, los autores comentan que la orientación dirigida a la medición de los resultados ganó fuerza en los años 1970 y en la década siguiente. Los gobiernos conservadores de este período —en especial en Inglaterra, con Thatcher, y en EEUU, con Reagan— imprimen reformas en el sector público, que fueron conocidas como el movimiento de la Nueva Administración Pública, preconizando la adopción de instrumentos de control y gestión de resultados del sector privado^{iv}.

Sin entrar en la discusión acerca de los excesos, limitaciones y aspectos meritorios de dicho movimiento —debate ya emprendido en el caso brasileño por varios autores, con mucho más competencia, autoridad y espacio, que el disponible para este texto—, el legado de preocupación con la transparencia y responsabilización pública creó oportunidades para el perfeccionamiento de la gestión de programas, sea para la estructuración de sistemas de indicadores de monitoreo, sea para la realización de las investigaciones de evaluación. En este tópico del texto, se discute el primero de los instrumentos de gestión: los sistemas de monitoreo.

El modelo idealizado de “ciclo de vida” de políticas y programas presentado anteriormente es una referencia conceptual interesante para ilustrar cómo los sistemas de indicadores de monitoreo pueden ser estructurados y cómo las investigaciones de evaluación pueden ser especificadas de manera que se potencialice su empleo en la gestión de los programas. La propuesta básica de esta sección es mostrar la importancia de disponer de indicadores relevantes y periódicamente actualizados para dar seguimiento a las actividades y a la producción de servicios de los programas, que permitan corregir desvíos, reprogramar actividades o inclusive especificar investigaciones de evaluación para entender por qué determinados procesos no están encaminándose en el sentido idealizado originalmente.

Como fue discutido en el texto anterior^v, las actividades de cada etapa del ciclo se apoyan en un conjunto específico de indicadores. En la definición de la agenda, los indicadores son recursos valiosos para dimensionar los problemas sociales, sirviendo como instrumentos de *advocacy* y presión de demandas sociales no satisfechas. Indicadores producidos por las instituciones oficiales de estadísticas —sobre todo los computados a partir de censos demográficos y encuestas muestrales regulares— se prestan bien para este papel, por la legitimidad que gozan ante diferentes públicos. En Brasil, el reporte “Síntesis de Indicadores Sociales”, publicado anualmente por el IBGE³, es una referencia importante en este sentido, provocando

³ IBGE : Instituto Brasileiro de Geografia e Estatística es la agencia estadística nacional de Brasil.

gran interés en los medios de comunicación cuando se divulgó^{vi}. El Índice de Desarrollo Humano (IDH) computado para países, municipios, mujeres y negros es otro buen ejemplo de cómo una medida simple puede movilizar los vehículos de comunicación y recolocar anualmente el debate sobre desigualdad social y los impactos de las políticas públicas en los medios de comunicación y en las esferas de gobierno^{vii}.

En la elaboración de los diagnósticos para la formulación de programas, los indicadores son imprescindibles para calificar a los públicos meta, localizarlos y retratarlos de modo tan amplio y detallado cuanto posible. Es necesario adecuar las intervenciones a las características y necesidades de los demandantes de los programas. De modo general, en la elaboración de diagnósticos socioeconómicos propositivos para programas públicos se emplean indicadores de varias áreas temáticas analíticas o de actuación gubernamental. Para desarrollar programas en el campo de la educación básica, por ejemplo, es necesario conocer no apenas las condiciones de acceso a la escuela, la infraestructura escolar, el desempeño de los alumnos, sino también aspectos relacionados a las condiciones de vida de los estudiantes, como vivienda, nivel de pobreza y recurso familiar, acceso a servicios de salud, escolaridad de los padres, factores que ciertamente pueden afectar o potencializar las acciones programáticas específicas^{viii}.

Indicadores contruidos a partir de los censos demográficos son particularmente útiles en esta fase, por la amplitud del enfoque temático investigado y por la posibilidad de desagregación territorial o por grupos sociodemográficos específicos. En Brasil, el Censo Demográfico 2010 potencializa aún más estas dos características —profundización y desagregabilidad—, constituyéndose en un marco importante para reevaluación de las demandas sociales de la población brasileña, en este momento de ampliación de la estructura de protección social en el país.

En la selección de alternativas programáticas idealizadas para atender a la cuestión pública colocada en la agenda y elegir prioridades de intervención, es necesario disponer de indicadores que operativicen los criterios técnicos y políticos definidos. Indicadores sintéticos, como el citado Índice de Desarrollo Humano, Índice de Desarrollo de la Educación Básica, Índice de la Huella Humana (*Footprint index*) o las clasificaciones generadas por técnicas de análisis multicriterio pueden ser útiles para apoyar decisiones en esta fase del ciclo, como discutido en Scandar *et ál.* (2006) y Jannuzzi *et ál.* (2009)^{ix}.

Para dar seguimiento a la implementación de los programas y, posteriormente, para su gestión, cuando entren en plena operación, son necesarios indicadores que permitan seguir regularmente las acciones programadas, de la previsión y asignación del gasto para la producción de los servicios y para los eventuales resultados e impactos esperados durante la idealización de los programas. Los indicadores que estructuran procesos formales de monitoreo, según Coutinho (2001, p.11), son como “un proceso sistemático y continuo que, produciendo informaciones sintéticas y en tiempo eficaz, permite una rápida evaluación situacional y la intervención oportuna que confirma o corrige las acciones monitoreadas”.

Dispuestos en paneles o en sistemas informatizados, los indicadores de monitoreo deben ser actualizados con regularidad y tempestividad adecuada con la toma de decisiones. Idealmente, estos indicadores deben ser específicos y sensibles a las acciones programadas, a fin de hacerse útiles para inferir si el programa está siendo implantado conforme lo planificado o para permitir las correcciones de rumbo (aspectos que sólo podrán ser comprobados más precisamente por medio de investigaciones de evaluación específicas, durante o después de la implantación).

Así, estos paneles o sistemas de indicadores de monitoreo deben valerse de los datos continuamente actualizados en los catastros y registros administrados por los gestores y operadores involucrados en los programas. Deben estar interconectados a los sistemas informatizados de gestión del programa, en el cual son registradas las atenciones, informaciones de los agentes que lo operan, características de los beneficiarios, procesos intermedios, que producirán los efectos idealizados por el programa. A diferencia de lo que sucede en las investigaciones de evaluación, en la estructuración de sistemas de monitoreo no se prevén levantamientos primarios de datos. Puede ser necesario crear rutinas de recolección de datos que operen fuera de los círculos normales de producción de servicios de los programas, pero es necesario hacer esfuerzos para aprovechar las informaciones generadas en el ámbito de operación cotidiana de los programas. La creación de una rutina más para registro de información, por el beneficiario del programa o por el agente involucrado en la implementación, puede implicar atrasos indeseados y, peor aún, declaraciones mal respondidas que acaban siendo inútiles para su finalidad original.

Se debe observar que, por lo general, las informaciones recopiladas y enviadas para la actualización de los registros de seguimiento del Plan Plurianual, en el Ministerio de Planificación, Presupuesto y Gestión, son solamente la punta del iceberg de datos almacenados en planillas electrónicas y gavetas en las oficinas de la administración pública. Además, es curioso que, en las medianas y altas gerencias del sector público —cabildos, secretarías de estados o ministerios—, se reclame de falta de información para la gestión en medio de tanta información continuamente producida en las unidades de prestación de servicios públicos (escuelas, centros de salud, comisarías, centros de referencia de la asistencia social, puestos de intermediación de mano de obra, agencias del INSS, etc.). Para superar esta paradoja de la "escasez en la abundancia", es necesario perfeccionar los procesos de gestión de la información en las oficinas donde se planifican y coordinan las políticas y programas públicos. Los flujos de información, los procesos de tratamiento, validación, clasificación y, por supuesto, de almacenamiento necesitan ser pensados de forma articulada, valiéndose de aplicativos y herramientas de integración de datos (y no de los costosos sistemas —en planificación, tiempo y dinero— informatizados de cuarta, quinta o enésima generación propuestos por consultorías con experiencia en automatización de procesos en la iniciativa privada, mas poco conocedoras de la complejidad operativa de la gestión de programas públicos).

Es posible que algunas informaciones producidas por las agencias estadísticas —en el caso brasileño el IBGE, Datasus, Inep, entre otros^x puedan ser útiles para la construcción de indicadores de monitoreo de programas, sobre todo de naturaleza universal o con población meta numerosa. Por lo general, estas fuentes proveen informaciones para evaluaciones de políticas o para un conjunto de políticas y macroacciones gubernamentales —más adecuadas para la composición de mapas estratégicos de la acción gubernamental— y no para los propósitos de monitoreo de programas específicos, en función de la escala territorial de desagregación de los datos o del tiempo que llevan para ser producidas. Este cuadro puede cambiar para mejor con las transformaciones anunciadas en las encuestas muestrales del IBGE para la década de 2010, con mayor integración conceptual, mayor posibilidad de incorporación de temas suplementarios, ampliación y ensanchamiento de la muestra en el interior del país. No sólo se podrá disponer de un enfoque más amplio de estadísticas e indicadores sociales divulgados más regularmente a lo largo del año, como también referidos para dominios territoriales y grupos sociodemográficos mucho más diversos. A los indicadores mensuales de empleo, hoy restrictos a las seis principales regiones metropolitanas, se sumarán muchos otros indicadores —educación, vivienda, ingresos, etc. — divulgados en base trimestral (semestral o anual) para todos los estados brasileños. Mediante el empleo de métodos estadísticos avanzados, usando información combinada de estas encuestas con otras fuentes de datos, investigadores de centros de análisis de políticas públicas y universidades podrán estimar indicadores sociales con buen grado de precisión para dominios territoriales aún más específicos^{xi}.

De cualquier forma, las informaciones estadísticas más generales son útiles para que se puedan disponer de algunos indicadores de contexto socioeconómico en el ámbito del sistema de monitoreo. Al final, todo sistema abierto, como son los programas públicos, está sujeto a los efectos de factores externos, que pueden potencializar o atenuar resultados. Sistemas de indicadores de monitoreo de programas de calificación profesional, por ejemplo, deben disponer de indicadores de mercado de trabajo y de producción económica como informaciones de contexto, por los impactos que una coyuntura económica menos o más favorable puede ocasionar en la operación del programa.

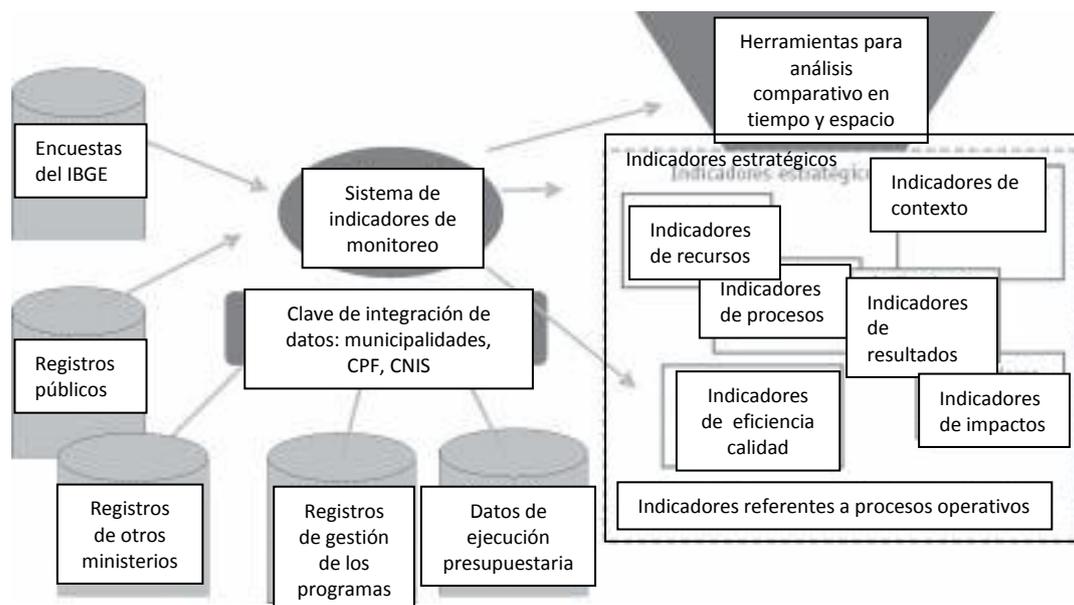
Como alternativa o complemento a las encuestas estadísticas oficiales, es posible construir indicadores de contexto o inclusive de monitoreo de programas, a partir de los registros administrativos de programas de gran cobertura poblacional como, en el caso brasileño, el Catastro Único de Programas Sociales del Ministerio de Desarrollo Social y Combate al Hambre; la Relación Anual de Informaciones Sociales y el Catastro General de Empleados y Desempleados del Ministerio de Trabajo; y el Catastro Nacional de Informaciones Sociales del INSS/Ministerio de Previsión Social. Naturalmente, los registros de proveimiento y ejecución presupuestaria del Sistema Integrado de Administración Financiera del Gobierno Federal (SIAFI) también son importantes fuentes de información para la construcción de indicadores de monitoreo (sobre todo porque permiten computar indicadores con regularidad de transferencia de recursos, dimensión crucial para programas que incluyen diversas acciones

intermediarias para viabilizar la consecución de las actividades más finalísticas). Vale registrar que los ministerios responsables por la gestión de estos registros han realizado esfuerzos importantes con el objetivo de poner a disposición sus informaciones para investigadores y público en general^{xii}.

Un buen sistema de indicadores de monitoreo no está necesariamente compuesto de una gran cantidad de información, sino de un sistema donde la información fue seleccionada de diferentes fuentes y está organizada de forma sintetizada y más adecuada para el uso analítico por los diferentes gestores (Diagrama 2). Es necesario encontrar un punto de equilibrio entre el "caos informacional", potencialmente generado por la estructuración de sistemas de monitoreo contruidos de abajo hacia arriba (donde participan inicialmente técnicos y gestores de la base y después de niveles tácticos y más estratégicos), y la pobreza analítica de las propuestas desarrolladas de arriba hacia abajo. Un sistema de indicadores de monitoreo no es un sistema de gestión operativa del programa, que suministra acceso a los incontables registros diarios e individuales de operación de convenios, prestación de servicios, recursos transferidos, proyectos y actividades concluidas. Un sistema de monitoreo se vale del (los) sistema(s) de gestión de los programas para buscar informaciones, integrarlas según unidades de referencia común (municipio, escuela, etc.), sintetizarlas en indicadores y conferirles un significado analítico. Al presentar informaciones sintetizadas en la forma de indicadores, que pueden ser analizados en el tiempo, por regiones y población meta, o que pueden ser comparados con metas esperadas, los sistemas de monitoreo permiten al gestor evaluar si los diversos procesos e innumerables actividades bajo su coordinación están "sumándose" en el sentido preconizado. Un sistema de monitoreo no es un conjunto exhaustivo de medidas desarticuladas, sino una selección de indicadores de procesos y acciones más importantes.

Un sistema que no suministra acceso orientado a las centenas de indicadores disponibles probablemente no se preste al propósito de monitoreo (aunque pueda ser útil como base de datos para estudios evaluativos a *posteriori*). Tampoco se presta al monitoreo un sistema donde la información no está organizada según el nivel de relevancia operativa-estratégica del gestor usuario. La información esencial debe estar disponible para el gerente de procesos operativos básicos para el buen desempeño de las actividades de sus coordinados. Deben estar disponibles para el gestor más estratégico, indicadores que le permitan dar seguimiento a los macroprocesos según el modelo lógico del programa.

Diagrama 2: Integración de informaciones de diferentes fuentes en la estructuración de un sistema de indicadores de monitoreo



Para uno y para otro gestor los indicadores deben ser los pertinentes a su esfera de decisión, ajustados a la referencia temporal y territorial que les compete e interesa. Las nuevas herramientas de integración de datos permiten construir paneles de indicadores de forma "customizada", permitiendo, inclusive, el acceso a la información más detallada si así el gestor lo desea. Se puede construir paneles en estratos "explicativos", es decir, organizando indicadores según una estructura nodal, donde un primer conjunto reducido de indicadores estratégicos sea seguido de un segundo conjunto más amplio de indicadores más específicos, que ayuden a entender el comportamiento y evolución de los primeros, y así por delante.

En realidad se trata de un sistema de monitoreo que reúne informaciones sintéticas o claves —para análisis de tendencias generales de las actividades estratégicas— e informaciones analíticas para un entendimiento más profundo de las tendencias observadas. La propuesta de seguimiento de las metas de inclusión social en los países de la Unión Europea sigue esta lógica de estructuración, disponiendo los indicadores en tres paneles articulados (ATKINSON *et ál.* 2005), como se ve en Diagrama 3.

Diagrama 3: Organización de Paneles de Indicadores en SM&E



Esta propuesta de organización de indicadores viene acompañada de algunas premisas básicas para seleccionarlos, que es oportuno rescatar, puesto que pueden ser útiles en otras aplicaciones (Cuadro 1). Vale también registrar que esta selección debería orientarse también por el análisis de la adherencia de los indicadores a las propiedades de relevancia social, validez de constructo, confiabilidad, periodicidad, sensibilidad a los cambios, especificidades de las acciones programadas, como discutido en Jannuzzi (2005).

Cuadro 1: Premisas para seleccionar indicadores de monitoreo de la inclusión / exclusión social en la Unión Europea

- El conjunto de indicadores no puede pretenderse exhaustivo y debe ser equilibrado entre las dimensiones de la exclusión social (salud, educación, vivienda, etc.). Un conjunto muy amplio de indicadores lleva a la pérdida de objetividad, pérdida de transparencia y credibilidad.
- Los indicadores deben tener una interpretación normativa claramente definida (para monitorear la exclusión social, la tasa de desempleo cumple dicho requisito; un indicador de productividad del trabajo no).
- Los indicadores deben ser mutuamente consistentes, es decir, no deben sugerir tendencias inconsistentes (indicadores de desigualdad como el índice de Gini y la proporción de masa salarial apropiada pueden tener comportamientos diferentes a lo largo del tiempo, ya que miden aspectos distributivos diferentes).
- Los indicadores deben ser inteligibles y accesibles a toda la sociedad. Son preferibles medidas simples, de fácil entendimiento. Se debe resistir a las simplificaciones indebidas (indicadores sintéticos).

El documento "Guía metodológica para la construcción de indicadores del PPA" (MPO, 2009), elaborado por el Ministerio de Planificación brasileño, es otra referencia importante para especificación de sistemas de indicadores de monitoreo de programas, por la extensa revisión bibliográfica utilizada en su elaboración, por la preocupación de incorporar aspectos de

propuestas desarrolladas en otros estados y contextos del sector público brasileño, por la experiencia de más de diez años en la elaboración de planes plurianuales en el gobierno federal, y por la preocupación con la claridad en la exposición de conceptos, etapas y fuentes de datos.

Si el programa fue especificado según las buenas prácticas y técnicas de planificación de proyectos, debe haber un diseño lógico de encadenamiento de actividades y etapas. Este sistema debe conseguir ofrecer evidencias sobre la ejecución del gasto, producción, eficiencia y calidad de los servicios, consumo y usufructo por parte de la población meta y, si es posible, anticipar dimensiones impactadas por el programa. Es decir, un buen sistema de monitoreo debe suministrar indicadores de insumo, proceso, resultado y posibles impactos del programa. Ejemplificando, este sistema debe permitir monitorear simultáneamente el dispendio realizado por algún tipo de unidad operativa prestadora de servicios o subproyecto; el uso de los recursos humanos, financieros y físicos; la generación de productos y la percepción de los efectos generados por los programas. En Rezende y Jannuzzi (2008) se presenta un ejemplo de panel de indicadores de monitoreo del Plan de Desarrollo de la Educación (PDE), estructurado en la lógica insumo-proceso-resultado-impacto. En la dimensión insumo se dio prioridad a los indicadores dirigidos al financiamiento de la educación; para la dimensión proceso fueron definidos indicadores que pudiesen ser producidos con cierta regularidad y que estuviesen relacionados al proceso de enseñanza-aprendizaje, como porcentual de docentes con nivel superior, porcentual de alumnos atendidos por turno integral, entre otros; como medida de resultados, los indicadores de aprovechamiento de la Prueba Brasil, desfase edad-grado y tasa de abandono; como impactos potenciales, el ingreso de jóvenes en la enseñanza superior^{xiii}.

La recopilación de indicadores y de programas organizada por la Secretaría de Evaluación y Gestión de la Información del Ministerio de Desarrollo Social y Combate al Hambre (DIC-VIP) es un producto útil para servir como referencia para la estructuración de sistemas de monitoreo de programas en esta perspectiva procesual, no apenas en este ámbito, sino para otros sectores de la administración pública^{xiv}.

En general, hay mayor disponibilidad de indicadores de procesos, que reflejan los esfuerzos y productos generados en las acciones programadas, para los cuales hay registros formales y mayor control operativo por los gestores. Aun así, en algunos casos es posible que se pueda disponer, en el panel de monitoreo de indicadores de resultados e impactos junto a la población meta de los programas, buscando informaciones en los registros y registros públicos citados anteriormente. En el caso del sistema de indicadores de monitoreo de programas de calificación profesional, ilustrado hace poco, sería posible disponer de alguna medida de impacto del programa, con la integración de informaciones catastrales de los trabajadores calificados con los datos de la RAIS o inclusive del CadÚnico, mediante el uso del número de identificación del trabajador o del Catastro de Persona Física (CPF) —de hecho, procedimientos análogos son usados corrientemente por el Ministerio de Desarrollo Social con fines de fiscalización en la concesión de beneficios del programa Bolsa Familia para familias registradas en el programa.

Con el avance de la informatización en el sector público brasileño es posible, también obtener informaciones referidas a unidades de prestación de servicios muy específicas (en pruebas, a nivel de escuelas, hospitales, puestos policiales, etc.), con buena periodicidad de actualización (frecuencia escolar mensual, atención ambulatoria semanal, etc.), lo que permitiría la construcción de indicadores de monitoreo referidos al contexto de vivencia de los beneficiarios de programas y al tiempo adecuado de la toma de decisiones. De hecho, proyectos sociales de alcance local han recurrido a la búsqueda periódica de informaciones en estos puestos de prestación de servicios públicos como estrategia de monitorear resultados e impactos de su acción. Centros de promoción de cursos de artesanía, actividades culturales y deportivas dirigidos a la reintegración social de jóvenes en comunidades muy violentas, proyectos que vinieron a surgir con frecuencia en los últimos años, por la acción directa de cabildos u organizaciones filantrópicas, pueden tener sus resultados e impactos inferidos por el eventual aumento de las tasas de frecuencia para la escuela, disminución de las atenciones ambulatorias por heridas o casos policiales involucrando jóvenes, entre otras informaciones recolectadas localmente.

Por supuesto, en los dos casos aquí ejemplificados —programa de calificación profesional y proyecto social de reintegración social de jóvenes— los indicadores citados pueden estar siendo afectados por otros factores — coyuntura más favorable del mercado de trabajo, en el primer caso, control policial más ostensivo, en el segundo— y no propiamente por la excelencia del programa o proyecto. No serían exactamente indicadores de impacto, pero quizás indicaciones potenciales de impacto, que para una efectiva atribución o vinculación causal con el programa o proyecto requerirían una investigación de evaluación específica. Diferentemente de estas últimas, donde la investigación de la atribución de un efecto a un programa puede ser una cuestión a ser evaluada, para inferencia *a posteriori*, en sistemas de monitoreo se asume, *a priori*, creyendo en la validez del diseño de implementación, la vinculación de acciones (y de sus indicadores) con los resultados (en la forma como pueden ser medidos por las informaciones generadas por el programa y otras fuentes secundarias de bajo costo de adquisición). (MCDAVID; HAWTHORN 2006).

Indicadores de eficiencia, productividad en la prestación de los servicios y de calidad de productos disponibles, como *proxies* de eficiencia y eficacia de los programas, también deberían constar de la matriz de indicadores de monitoreo. En programas de calificación profesional, por ejemplo, con base en los registros de operación de los propios programas es posible computar indicadores como costo de formación por alumno, número promedio de alumnos formados por entidad acreditada, tasa de frecuencia o abandono de calificados.

En Brasil, hay cinco sistemas de macro-seguimiento de programas- MONIB – Painel de Indicadores do Brasil Sem Miséria; Siga Brasil del Senado, Matriz de Información Social-MIS de la Sagi/MDS, SIM-Trabalho de la Secretaría de Empleo y Relaciones del Trabajo de Sao Paulo y el programa Agenda 2012 del Ayuntamiento de Sao Paulo, que pueden ilustrar algunos de los aspectos estructurales aquí presentados^{xv}.

Talvez la experiencia más completa en términos de las propuestas presentadas, es la que se desarrolla en la Secretaria de MDS, con Painel de Indicadores do Brasil Sem Miséria – MONIB – y Painel de Conjuntura, que organizan paneles de indicadores en diferentes niveles e temáticas, con diferentes formas de presentación y funcionalidades.

Las otras cuatro experiencias relacionadas no son propiamente sistemas de indicadores de monitoreo para gestores de programas, sino herramientas para consulta y seguimiento de programas públicos por parte del ciudadano, ya que están franqueados a los usuarios de la Web. El programa Siga Brasil permite acceso a datos e indicadores históricos sobre la ejecución presupuestaria del gobierno federal, a nivel de programas, para estados y municipios. La aplicación en línea MIS reúne un conjunto amplio de datos de contexto e indicadores de programas federales, de interés a gestores e investigadores del área de asistencia social. El programa SIM-Trabalho es un sistema dirigido a la elaboración de diagnósticos para especificación de planes locales de calificación profesional permitiendo también acceso a informaciones sobre la ejecución de programas de intermediación de mano de obra y microcrédito. La Agenda 2012 es un portal de comunicación del Ayuntamiento de Sao Paulo que trata de la evolución de las metas de gestión, en las varias secretarías, por distritos y subdistritos, con actualización semestral. Fue creada para responder a las exigencias de la Enmienda 30 a la Ley Orgánica del Municipio, aprobada por la Cámara Municipal, al finales de 2008. Esta iniciativa tendrá, ciertamente, repercusiones en otras localidades del país.

Como ya fue señalado, un buen sistema de monitoreo debe permitir el análisis comparativo de los indicadores a lo largo del tiempo y para distintas unidades de prestación de los servicios o atención del programa, así como en relación a estándares normativos de referencia. El análisis de la evolución de los indicadores, de sus avances, retrocesos y permanencias permite identificar fallas operativas del programa, aunque no la explicación de las mismas, en general. Esta es una diferencia básica entre los sistemas de monitoreo y las investigaciones de evaluación: los primeros, al asumir que existe una lógica de encadenamiento de actividades, delinean comportamientos esperados para los indicadores; en las investigaciones de evaluación, alertados por los desvíos no esperados en la evolución de los indicadores, se buscan las explicaciones valiéndose de los métodos y técnicas de la investigación social (MCDAVID; HAWTHORN, 2006).

En el ejemplo anterior, si un gestor necesita entender por qué la tasa de abandono de alumnos en cursos de calificación en una dada localidad o en una institución está aumentando quizás consiga alguna interpretación para la situación en su propio sistema de monitoreo, si fuera un sistema que dispone de informaciones analíticas, además de sintéticas. Puede ser que haya indicadores de contexto que muestren un fuerte aumento de las oportunidades en el mercado de trabajo regional (de allí la evasión). La extracción de informaciones de este sistema hacia un paquete estadístico o aplicativo de extracción de datos puede permitir al gestor refinar su hipótesis, si dispone de conocimiento o consultoría especializada en el uso de las herramientas^{xvi}.

O quizás no tenga ninguna pista adicional y busque una explicación con el dirigente de la institución o gestor responsable por el área. Es lo que Worthern *et ál.* (2004, p.38) clasifican como evaluación informal, que ocurre "siempre que una persona opta por una entre varias alternativas existentes, sin antes haber recolectado evidencia formal del mérito relativo de estas alternativas". Aunque no sea pautada en procedimientos sistemáticos, estas evaluaciones no siempre "ocurren en el vacío".

La experiencia, el instinto, la generalización y el razonamiento pueden influenciar el resultado de las evaluaciones informales, y cualquiera de estos factores, o todos, pueden ser la base de buenos juzgamientos (WORTHERN *et ál.* 2004, p.38).

Si su conjunto de indicadores de monitoreo no le permite llegar a explicaciones satisfactorias, si sus estrategias de visita o contacto informal no le agregan evidencias convincentes, o si el proceso de evasión gana dimensiones preocupantes en algunas áreas y no en otras, quizás sea el momento de que el gestor piense en encomendar una investigación de evaluación.

Las investigaciones de evaluación en el ciclo de políticas y programas

El hecho de denominar 'evaluación' a la fase posterior a la implementación de programas en el ciclo de políticas y programas crea una distinción artificial entre las actividades de monitoreo ejercidas con base en los sistemas de indicadores descritos en el tópico anterior y las actividades de realización de investigaciones de evaluación, que pueden ser realizadas en cualquier momento del ciclo.

En realidad, monitoreo y evaluación de programas son términos para designar procedimientos técnicos formales de seguimiento de programas, enfocados al análisis de la eficiencia, eficacia y efectividad de los mismos, para perfeccionar la acción pública (COHEN; FRANCO 1999). Monitoreo y evaluación son procesos analíticos orgánicamente articulados, complementándose en el tiempo, con el propósito de subvencionar al gestor público con informaciones más sintéticas y tempestivas sobre la operación del programa —resumidas en paneles o sistemas de indicadores de monitoreo— e informaciones más analíticas sobre su funcionamiento, obtenidas en las investigaciones de evaluación.

Como bien conceptúa el documento del Tribunal de Contas da União brasileño:

El monitoreo y evaluación de los programas de gobierno son herramientas esenciales para la buena práctica gerencial. La evaluación es un procedimiento que debe ocurrir en todas las etapas permitiendo al gestor federal el seguimiento de las acciones y su revisión y reconducción cuando sea necesario. Mientras que el monitoreo es una actividad gerencial interna, que se realiza durante el período de ejecución y operación, la evaluación puede ser realizada antes o durante la implementación, al concluir una etapa o proyecto como un todo, o inclusive algún tiempo después, debiendo preocuparse con el impacto

provocado por la intervención pública en sus beneficiarios (TCU, 2006, p. 75).

Las diferencias y especificidad de la función y de las actividades de monitoreo y evaluación pueden ser ilustradas a partir de un ejemplo más simple: para medir y acompañar la evolución de la fiebre de un niño podemos usar la palma de la mano o el termómetro – **instrumentos de monitoreo** – que nos proveen información crucial, rápida y barata para una toma tempestiva de decisiones de ofrecer una pastilla contra la fiebre o entonces llevarlo para un centro médico para una **evaluación** clínica más detallada de las razones del estado febril del infante, más costosa y que requiere más tiempo.

Las tipologías clásicas usadas para clasificar las investigaciones de evaluación también acaban favoreciendo alguna confusión, como por ejemplo considerar monitoreo como evaluación de proceso. Esta última es realizada cuando se requiere evaluar las actividades de implementación de los programas, la extensión de cobertura de la población meta atendida, la calidad de los servicios ofrecidos o entender las dificultades que están “trabando” la efectiva operación del programa, en contraposición a la investigación de evaluación de resultados e impactos realizado idealmente después de haber sido resueltos los problemas de gestión, cuando se requiere en la fase de evaluación del ciclo, una respuesta cabal a las preguntas: ¿el problema social que originó el programa fue determinado? ¿Cuál es la contribución del programa para este caso? ¿El programa debe continuar, ser expandido o ser descontinuado? (CARVALHO, 2003).

Cohen y Franco (1999) las clasifican como evaluación formativa y evaluación sumativa, respectivamente. Estas últimas serían evaluaciones del tipo *ex-post*, que por su vez serían distintas de las evaluaciones *ex-ante*, destinadas a estudios de necesidad y viabilidad de creación de un programa. Worthern *et ál.* (2004) rescatan otra clasificación clásica, diferenciando las investigaciones de evaluación según sus protagonistas: evaluación interna, conducida por el equipo del programa; la evaluación externa, realizada por el equipo de consultores externos contratados; la evaluación mixta, que prevé la interacción de los dos grupos; y la evaluación participativa, construida junto con los beneficiarios de los programas (que también contribuyen para su planificación).

Cada modalidad tiene ventajas y limitaciones: si la evaluación externa parece permitir una visión exterior no sesgada y no comprometida con la continuidad del programa, confiriendo mayor exención al análisis (suponiendo que los consultores son absolutamente profesionales, no preocupados con las oportunidades de contratación en una próxima demanda de evaluación); la evaluación interna, la mixta y la participativa parecen permitir aportes más circunstanciados en el análisis de los procesos y resultados, además de facilitar la incorporación de mejoras en la gestión del programa —objetivo final de la evaluación—. Evaluaciones internas pueden ser rápidas y baratas, al contrario de otras modalidades. Evaluaciones participativas pueden viabilizar perfeccionamientos más efectivos.

Una tipología de investigaciones evaluativas particularmente interesante, en virtud de su adherencia al ciclo de políticas y programas ya expuesto, es propuesta por Rossi *et ál.* (2004) (Diagrama 4). Para estos autores, los estudios evaluativos deberían componer un programa de investigación completo, siguiendo la lógica implícita del "ciclo de vida" de los programas: antes de nada, sería necesario analizar la pertinencia del programa en la agenda de políticas públicas (*Needs assessment*), investigando si el programa responde a una demanda social efectivamente existente —lo que propondría la evaluación *ex-ante* de Cohen y Franco (1999).

Si el programa tuviera razón de existir porque procura responder a una demanda pública de hecho, y no a una necesidad de autopreservación de una estructura organizativa creada en algún momento del pasado, debería investigarse la teoría o modelo de intervención donde el programa se asienta (*Design assessment*). Al final, ¿Este es el mejor diseño lógico de intervención, en el cuadro de relaciones institucionales que prevalece entre las esferas de gobierno? ¿Frente a las distintas capacidades de gestión y control social en otros países? Respondidas estas preguntas, haría sentido entonces investigar el proceso de implementación del programa (*Program process assessment*), y si se constata la inexistencia de graves problemas referentes a la entrega de los servicios a la población meta, haría sentido evaluar impactos y resultados del programa (*Impact assessment*). Al final, habiendo constatado que el programa es socialmente justificable, cuenta con un diseño de implementación adecuado, con procedimientos operativos bien establecidos y con resultados efectivamente comprobados, se pasaría a la evaluación de la eficiencia o costo-efectividad (*Efficiency assessment*), para responder si el costo de la operación del programa se legitima por los efectos obtenidos (en términos de *outputs* monetizables o en indicadores sociales efectivos) y si es replicable para otras escalas de operación.

La propuesta de Rossi *et ál.* (2006) es ciertamente un avance conceptual en la forma de entender los diferentes focos de los estudios evaluativos, más allá de las dicotomías evaluación de procesos/evaluación de impactos, evaluación *ex-ante*/evaluación *ex-post*, etc. Al proponer énfasis evaluativos para cada etapa del "ciclo de vida" del programa, los autores ofrecen un plan de evaluación amplio, orientando la especificación de las investigaciones de evaluación según una lógica de abordaje que tendería a garantizar, *a priori*, resultados relevantes y útiles al final del proceso. Esta propuesta ayudaría a definir mejor la pregunta crucial para el éxito de una investigación de evaluación de programas: qué aspectos del programa deben ser evaluados y cuáles deben ser abordados posteriormente, en función del nivel de madurez del programa, para maximizar las oportunidades de obtenerse informaciones útiles y reveladoras para su perfeccionamiento en el momento.

When developing the questions around which the plan for an evaluation will resolve, therefore, it is best for the evaluator to start at the bottom of the evaluation hierarchy [evaluación de la necesidad del programa] and consider first what is known and needs to be known about the most fundamental issues. When the assumptions that can be safely made are identified and the questions that must be answered are defined, then it is appropriate to move to the next level of hierarchy [...]

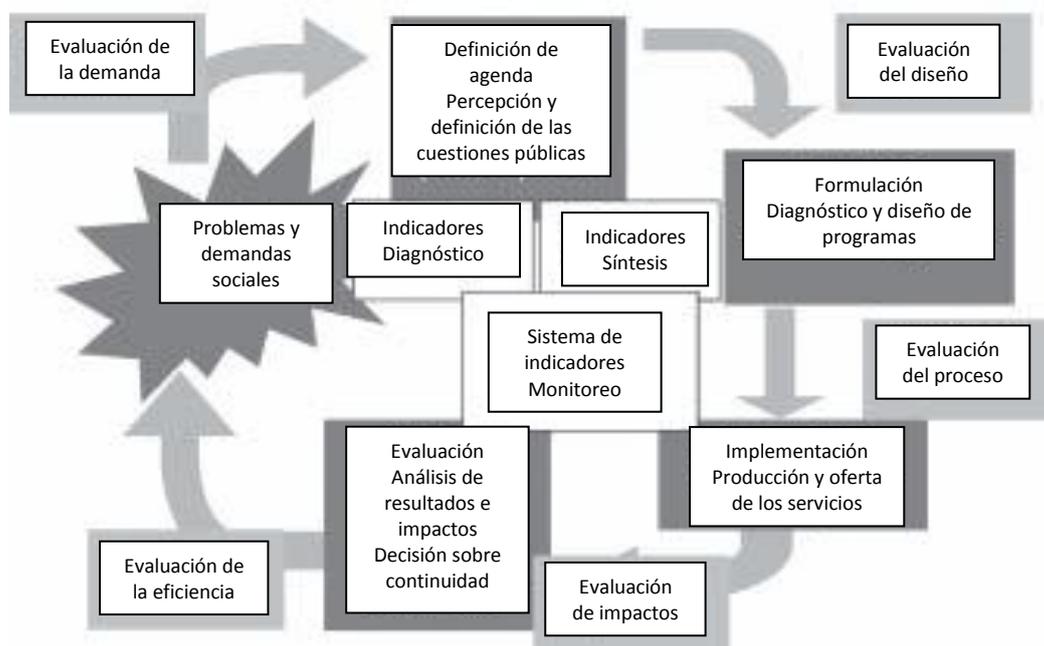
By keeping in mind the logical interdependencies between the levels in the evaluation hierarchy and the corresponding evaluation building blocks [técnicas e instrumentos apropiados de la investigación social], the evaluator can focus the evaluation on the questions most appropriate to the program situation (ROSSI et ál.2004, p.81).

Al final, no correspondería especificar una investigación de evaluación de resultados e impactos si hay dudas en relación al nivel de madurez de los procesos y acciones previstas en la implementación del programa o si existen sospechas de que el diseño de implementación padece de supuestos no sostenibles en su estructura lógica de intervención. No correspondería proceder a investigaciones complejas de análisis de costo-efectividad si no hay certeza de que los resultados e impactos generados por el programa son significativos. En fin, evidencias de baja efectividad del programa podrían ser consecuencias menos de fallas en el diseño o de la forma como fue implementado y más de la propia inadecuación de la existencia del programa, como parece ser el caso de los programas dirigidos a ofrecer la experiencia del primer empleo para jóvenes (MADEIRA, 2004).

La jerarquía de focos evaluativos de Rossi *et ál.* (2006) no implica que el gestor sólo pueda disponer de informaciones acerca de la eficiencia, de la calidad de los servicios, de la cobertura de la atención del público meta de los programas en las fases más adelantadas del "ciclo de vida" del programa. Al final, se presupone que hay un sistema de indicadores de monitoreo ya estructurado que reúne informaciones indicativas en este sentido. Inclusive, puede ser necesario encomendar investigaciones de evaluación de resultados e impactos o estudios más profundos de evaluación de la eficiencia en etapas aún incipientes de la implementación del programa, para responder a demandas externas de evaluación. Lo que los autores defienden es que se organice un plan consistente de evaluación en todas las etapas del "ciclo de vida" del programa, y que se resista a la tentación de "poner la carreta delante de los bueyes", encomendando investigaciones de evaluación de forma precoz y desarticulada. Es necesario ser diligente con el riesgo de producción de resultados irrelevantes o, peor, con la producción de resultados precipitados en pretensa legitimidad científica.

Cada uno de estos tipos de estudios evaluativos requiere de un abordaje metodológico y/o disciplinario diferente. Estudios de viabilidad o necesidad de programas pueden valerse de análisis documentales comparativos, de recurso para las técnicas de análisis económico o inclusive para los análisis institucionales típicos de la ciencia política. Análisis de diseños lógicos de implementación implican la revisión de documentos usados en la formulación del programa, empleo de las técnicas de planificación de proyectos y de comparación con otros modelos de intervención social similares en otros contextos y momentos.

Diagrama 4: Indicadores e investigaciones de evaluación en el ciclo de políticas y programas



La evaluación de procesos puede requerir el empleo de técnicas variadas de recolección de informaciones usadas en la investigación social. Evaluación de resultados e impactos puede requerir el empleo de entrevistas con agentes institucionales, grupos de discusión con beneficiarios de los programas, encuestas muestrales, delineaciones cuasi experimentales o estudios comparativos de casos.

La especificación del delineamiento metodológico adecuado con los objetivos de la investigación de evaluación pretendida es un aspecto crucial. Al final, la evaluación de programas es un emprendimiento técnico-científico de uso de métodos de la investigación social para establecer la situación, problemas y diferentes aspectos de la gestión de un programa público a lo largo de su "ciclo de vida", de su concepción al usufructo de los productos y servicios puestos a disposición, considerando el contexto organizativo y político donde se inserta, con la finalidad de informar las necesidades de perfeccionamiento de sus acciones, de manera que se contribuya, junto con otros programas, a mejorar las condiciones sociales de la población. Como emprendimiento técnico-científico, así como en la conducción de investigaciones académicas, la selección del método depende de los objetivos específicos anhelados por el evaluador. La búsqueda de datos secundarios, la observación y sus variantes —visita, observación participante, etc.—, las entrevistas —semiestructuradas o estructuradas—, los grupos de discusión, las encuestas de campo con cuestionarios —con muestras probabilísticas o intencionales— los diseños cuasi experimentales y no experimentales son algunas de las estrategias metodológicas más empleadas en las investigaciones de evaluación de programas en Brasil.

Cada técnica tiene sus características, ventajas y sus limitaciones. Los métodos cuantitativos, organizados bajo la égida del modelo hipotético-deductivo, así como los levantamientos muestrales y experimentos comprenden técnicas bastante estructuradas, destinadas a la investigación de problemas específicos, dirigidas al dimensionamiento de cantidades o intensidad de relaciones entre variables, suponiendo distanciamiento del investigador en relación al objeto investigado. Los métodos cualitativos, basados en el principio de la producción del conocimiento por la lógica inductiva, de lo particular para lo general como la observación, entrevistas y grupos de discusión, comprenden técnicas poco o semiestructuradas, para investigar exploratoriamente problemas complejos, presuponiendo la proximidad del investigador con el objeto evaluado.

Escapa al objetivo de este trabajo detallar prescripciones metodológicas acerca del empleo de una u otra técnica, dada la complejidad de cada una y por la buena disponibilidad de manuales de métodos y técnicas de la investigación social en Brasil, como los de Laville y Dionne (1999), Cano (2002), Richardson *et ál.* (2002) y Babbie (1999), entre otros. Vale incluir en esta lista, como un importante material de referencia para el delineamiento metodológico de investigaciones de evaluación, el documento organizado por la Secretaría de Evaluación y Gestión de la Información del Ministerio de Desarrollo Social y Combate al Hambre de Brasil (MDS, 2007). Son presentados 18 estudios, en su mayoría investigaciones de evaluación de programas del MDS en el área de Asistencia Social, Seguridad Alimentaria y Nutricional, e Ingresos de la Ciudadanía. En la exposición del estudio evaluativo de cada programa son presentadas las instituciones e investigadores participantes del estudio, el período de realización, los objetivos de la evaluación, los aspectos metodológicos acerca de las técnicas de recolección de datos empleadas, las características de la muestra y sujetos entrevistados y los instrumentos de recolección de datos (cuestionarios o guías de entrevistas). Las bases de datos de varias de estas investigaciones de evaluación fueron puestas a disposición en el sitio de la Secretaria e en el Consorcio de Informaciones Sociales de la Anpocs, permitiendo que investigadores y evaluadores busquen respuestas preliminares para preguntas de otros programas similares^{xvii}.

Este material evidencia en la práctica, cómo la naturaleza del programa, la etapa en que se encuentra, los objetivos de evaluación, la disponibilidad de tiempo y recursos acaban orientando las decisiones metodológicas acerca de instrumentos y técnicas de investigación, el tamaño, intencionalidad o no de las muestras de las investigaciones de evaluación encomendadas por la SAGI. Se ilustra, cómo las investigaciones de evaluación de proceso conducidas para varios programas procuran garantizar la obtención de informaciones a través de entrevistas semiestructuradas con agentes públicos en diferentes posiciones de la producción de los servicios (gestores municipales, técnicos en atención al público), en localidades intencionalmente seleccionadas (según el tiempo de implantación del programa, por ejemplo).

Grupos de discusión con equipos técnicos involucrados en los programas y sobre todo con sus beneficiarios, permiten obtener rápidamente informaciones cruciales para el perfeccionamiento de procedimientos y

acciones de los programas y también para captar dimensiones de impacto que no son fácilmente tangibles. Un ejemplo ilustrativo de esta técnica es descrito por Carvalho (2003), en la estrategia para evaluar los resultados e impactos del programa Joven Ciudadano. Este programa proporcionaba a los estudiantes de 16 a 21 años de la red pública provincial de enseñanza secundaria, la primera oportunidad de experiencia profesional en el mercado de trabajo.

Por tratarse de un público compuesto por jóvenes, la técnica utilizada fue la de grupos de discusión, aplicándose "dinámicas que estimulen y faciliten la manifestación de los participantes, como simulaciones y juegos, haciendo la investigación casi lúdica. Estas dinámicas ayudan a atenuar las dificultades como inhibición, timidez, desconfianza de determinados públicos. Es particularmente útil en el caso de jóvenes, segmento que generalmente se caracteriza por conversaciones lacónicas, en código, en especial ante adultos" [...] (CARVALHO 2003, p. 189).

Análisis estructural de textos es una técnica usada para interpretar relatos de entrevistas o de grupos de discusión de forma más estructurada, tratando palabras, frases, párrafos como entidades empíricas, permitiendo contabilizar frecuencias, recurrencias y relaciones entre ideas de un conjunto de textos y documentos (BOWEN & BOWEN 2008). En el campo de las investigaciones de evaluación de programas, en especial en las encuestas de satisfacción o evaluación de impacto junto a beneficiarios de programas sociales, la técnica puede ser utilizada como recurso analítico para sistematizar de forma más objetiva y estandarizada las manifestaciones, opiniones y críticas presentes en los discursos de los entrevistados, y como estrategia metodológica para garantizar mayor replicabilidad de las evaluaciones en otros contextos territoriales y temporales. Worthen *et ál.* (2004) estiman estas técnicas como avances metodológicos significativos en los enfoques cualitativos de evaluación.

No existe *a priori* un método universal más legítimo o con mayor "status científico" para toda y cualquier investigación de evaluación, como no existe un único método para las investigaciones académicas. La realidad social es compleja y los problemas sociales son multidimensionales y pueden – y deben– ser analizados según diferentes abordajes disciplinarios – sociología, economía, ciencias políticas, antropología, etc.- y diseños metodológicos – cualitativos, cuantitativos etc. Cada disciplina científica tiene sus abordajes más convencionales y legitimados. Además, inclusive dentro de un campo de conocimientos disciplinarios hay distintas comunidades de investigadores, que comparten sus propias visiones de mundo y paradigmas teórico-metodológicos.

Como fue señalado en un importante manual de investigación social americana acerca de la práctica de la investigación científica:

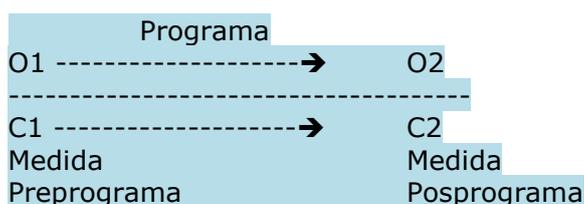
1. It seems to us futile to argue whether or not a certain design is "scientific"
[...] It is not a case odd scientific or not scientific, but rather one of good or less good design [...]
2. The proof of the hypotheses is never definitive [...]

3. There is no such thing as a single “correct” design. Different workers will come up with different designs favoring their own methodological and theoretical predispositions [...]
4. All research design represents a compromise dictated by the many practical considerations that go into social research. None of us operates except on limited time, money, and personnel budgets [...] A research design must be practical.
5. A research design is not a highly specific plan to be followed without deviation, but rather a series of guideposts to keep one headed in the right direction [...] (SUCHMAN in MILLER 1991, p. 98).

Así, es extraño que aún persista en ciertas comunidades el mito de que los lineamientos experimentales o cuasi experimentales se constituyen en los métodos más adecuados y legítimos científicamente para evaluación de impactos de programas sociales, cuyas dimensiones “impactadas” suelen ser muchas veces difícilmente tangibles por indicadores cuantitativos.

Estos diseños metodológicos de evaluación son inspirados en el formato clásico del experimento en las Ciencias Naturales. Dicho modelo tiene el objetivo de investigar la estructura e intensidad de causalidad entre una variable-efecto dependiente y sus factores determinantes. Para ello es preciso garantizar el control de la situación experimental en laboratorio y el empleo de grupos de tratamiento (beneficiarios de programas) y control (no beneficiarios) escogidos de forma aleatoria. En los diseños cuasi experimentales, una de las condiciones básicas que definen el experimento y que no es verificada, por lo general, es la designación aleatoria en grupos experimentales y de control. Como es ilustrado en el Cuadro 2, el diseño evalúa la evolución de los dos grupos —supuestamente idénticos en el comienzo del experimento— a lo largo del tiempo, y constatar si al final hay evidencias de que el primero mostró mejor *performance* que el segundo en una variable entendida como reveladora de los efectos del programa.

Cuadro 2: Diseño experimental clásico



O: Grupo de tratamiento (beneficiario del programa)

C: Grupo de control (similar al otro grupo, pero no es beneficiario)

Grupo de individuos O y C definidos por designación aleatoria

Si $O2 - O1 > C2 - C1$ o si $O1 = C1$ y $O2 > C2$ entonces el programa produce impacto

En realidad esta línea de investigación —experimental o cuasi experimental— es uno de los métodos usados para la evaluación de impacto, no necesariamente el mejor o más factible. En este tipo de diseño evaluativo, hay problemas éticos (¿cómo escoger y justificar quién será el beneficiario y quién quedará fuera del programa?), operativos (¿cómo evitar la evasión de los beneficiarios, cómo garantizar que los efectos medidos son

apenas del programa enfocado, en medio de un contexto creciente de intervenciones sociales?), metodológicos (¿el indicador empleado es la mejor medida para captar la dimensión impactada? Lo que se supone como dimensión impactada, por el diseño lógico del programa, ¿guarda un estrecho vínculo con las acciones y actividades del programa? ¿El impacto debe ser medido sobre los beneficiarios o junto a la comunidad a la que pertenecen?) y epistemológicos (¿cómo los esfuerzos de garantía de la validez interna del experimento conspiran contra la generalización de los resultados? ¿cómo garantizar que un programa bien evaluado en circunstancias tan artificiales pueda repetir el éxito en situaciones normales?) (WEISS 1972, MOHR 1995, CANO 2002).

Aunque todos estos problemas se pudiesen contornear, quedaría un problema práctico: si los efectos potenciales del programa, tal como medidos en una determinada variable, no fueren elevados —algo que el incrementalismo de Lindblom (2006) sugeriría—, las muestras de beneficiarios atendidos y del grupo de control tendrían que ser considerablemente grandes para que las pruebas estadísticas puedan ser aceptadas sin hesitación (ROSSI *et ál.* 2004).

La mitificación de este diseño en la evaluación de programas se debe, en alguna medida, al origen de los estudios evaluativos centrados en la investigación de programas en el área de educación y salud pública, como ya fue mencionado, donde dichos modelos pueden hacerse viables más concretamente por las condiciones de simulación de “laboratorio” en salas de clases o por la tradición de los ensayos de tratamiento clínico de enfermedades. La hegemonía circunstancial de los modelos cuantitativos importados de la investigación en ciencias naturales en el debate sobre la cientificidad de los métodos de investigación a ser enseñados y adoptados en la investigación social americana en los años 1960, momento de expansión de los estudios evaluativos en EEUU —como ya fue mencionado— también es otro factor explicativo. En este contexto, como sugieren Worthern *et ál.* (2004), el libro “*Experimental and quasi-experimental designs for research*”, escritos por los psicólogos Donald Campbell y Julian Stanley, publicado en 1966, acabó siendo adoptado como el manual de referencia metodológica de la comunidad de evaluadores que se estaba formando. Al final, a pesar de que las advertencias sobre las dificultades de replicación de las condiciones de control en laboratorio en el contexto de operación de los programas sociales, “la elegancia y la precisión del método experimental llevaron a la mayoría de los evaluadores de programas a verlo como ideal” (WORTHERN *et ál.* 2004, p. 116).

Las críticas en las décadas siguientes sobre aspectos éticos, factibilidad operativa y poder de generalización de los resultados de diseños experimentales —y sus aproximaciones cuasi experimentales- sea en la investigación académica, sea en la investigación de evaluación de programas, la incorporación de evaluadores provenientes de varias disciplinas de las ciencias sociales: antropólogos, sociólogos, comunicólogos, etc. y la formalización más rigurosa de enfoques de investigación más cualitativos, más adecuados a los problemas complejos y poco estructurados de la realidad social— acabaron por consolidar la percepción en la comunidad de evaluadores en los EEUU de que los estudios

evaluativos requieren cierto eclecticismo metodológico, integrando métodos cuantitativos y cualitativos.

En verdad, la mayoría de los evaluadores serios comenzó a entender que los enfoques cuantitativos y naturalistas tienen estándares metodológicos y rigor que son diferentes y no ausentes. La mayoría de los evaluadores pasó a aceptar la existencia de las múltiples realidades o por lo menos de las múltiples percepciones de la realidad. Con esta conciencia y una legitimidad mayor, la evaluación cualitativa surgió como alternativa real —o complemento— del enfoque cuantitativo tradicional (WORTHERN et ál. 2004, p. 117).

Aun así, los autores continúan alertando que el debate no está concluido, a medida que:

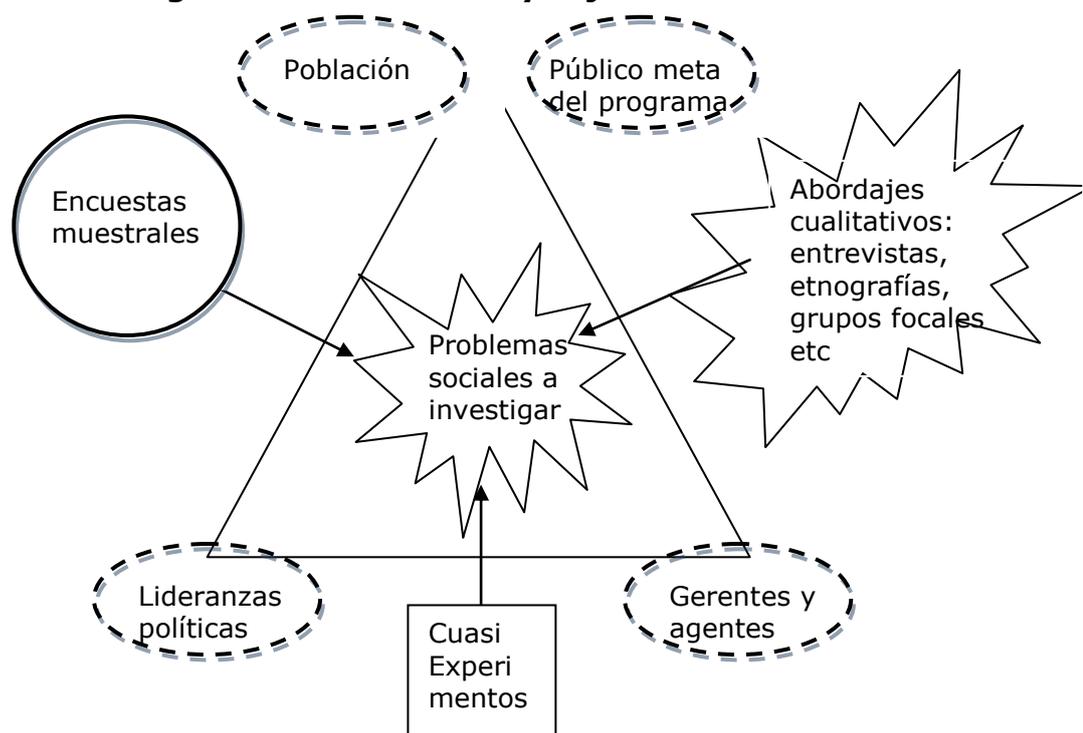
Aquellos que prefieren el uso exclusivo o mayoritario de métodos cuantitativos están, en su mayor parte, fastidiados con la aceptación de los estudios cualitativos (a pesar del hecho de que el trabajo cuantitativo aún mantiene su posición como enfoque dominante de la evaluación e investigación) [...] Estos críticos de la evaluación cualitativa se quejan con frecuencia de la subjetividad de muchos de los métodos y técnicas cualitativas, mostrando la preocupación de que la evaluación haya abandonado la objetividad en favor de una subjetividad ejercida inhábilmente (WORTHERN et ál. 2004, p.117).

Dada la complejidad operativa, los contextos de implementación, los diseños institucionales y diversidad de públicos meta de los programas públicos en Brasil y en países de Iberoamérica no se puede definir una técnica de investigación como siendo la mejor en cualquier situación. Muchos programas operan en contextos complejos, poco estructurados para enfoques cuantitativos (y mucho menos para diseños cuasi experimentales). En estas situaciones, enfoques metodológicos menos estructurados pueden obtener evidencias más relevantes y útiles para el perfeccionamiento de los programas. Es posible que permitan la estructuración de preguntas más específicas para investigar, para lo que los métodos cuantitativos pueden revelarse más útiles. Enfoques cuantitativos y cualitativos no son mutuamente excluyentes en un proyecto de investigación o evaluación. Son complementarios, compatibles y conectables.

Así, la estrategia evaluativa debe ajustarse a la complejidad del problema a ser investigado, y no lo contrario. Es el problema de evaluación que debe conducir la selección de los métodos, sujetos, instituciones a ser investigadas. No es la supuesta consistencia técnica de los métodos de evaluación la que debe condicionar la forma en que el programa deba ser conducido. El diseño o estrategia de implementación del programa debe ser aquello que mejor se ajusta a los condicionantes, restricciones e fortalezas existentes en el contexto de operación del programa.

Problemas sociales complejos requieren estrategias evaluativas igualmente complejas, requiriendo muchas veces la triangulación de metodologías y sujetos, o mejor, la mirada del objeto de estudio con técnicas diferentes de investigación o sujetos entrevistados (Diagrama 5).

Diagrama 5: “Problemas complejos, evaluaciones complejas” por medio de triangulación de métodos y sujetos



Pluralismo metodológico, enfoques evaluativos mixtos, triangulación de enfoques investigativos, complementariedad de técnicas son las perspectivas de la investigación de evaluación de programas compartidos por los autores de las dos principales referencias bibliográficas aquí citadas —Worthern *et ál.* (2004) y Rossi *et ál.* (2004). Rigor metodológico y maleabilidad técnica— con instrumentos de naturaleza más estructurada o menos estructurada, más cuantitativos o más cualitativos, cuasi experimentales o no, frente a la complejidad del objeto de estudio, éstas son las prescripciones generalizables para cualquier equipo —necesariamente multidisciplinario— que quiera enfrentar responsablemente la investigación aplicada a la evaluación de programas.

Por fin, se debe observar que las investigaciones de evaluación no implican necesariamente obtener datos a través de algunas de las técnicas anteriormente listadas. Los datos disponibles en los registros operativos del programa, el sistema de indicadores de monitoreo, las encuestas de los Institutos Nacionales de Estadística y datos de otros órganos pueden permitir realizar análisis preliminares a bajo costo y tiempo. El problema es que, por lo general, no hay la especificidad necesaria para responder a preguntas más concretas. Pero si es posible integrarlas por medio de claves – individuales o por municipios – se puede hacer análisis econométrico con cierto poder evaluativo.

Otras fuentes útiles de información secundaria son los reportes de evaluación de programas de órganos de Control Público, que en Brasil son el Tribunal de Contas da União (TCU) y de la Contraloría General de la Unión (CGU)^{xviii}. Los reportes de evaluación de programas del TCU,

elaborados desde 1998, son organizados en grandes tópicos, iniciándose con la exposición de los objetivos de la evaluación del programa (sección "Lo que fue evaluado"), de la justificativa para su realización ("Por qué fue evaluado"), de los aspectos metodológicos de la investigación evaluativa, explicando métodos y técnicas de recolección y análisis de los datos, las muestras y sujetos estudiados ("Cómo se desarrolló el trabajo) y una breve presentación del programa ("Histórico del programa"). Los resultados de la evaluación ("Lo que el TCU encontró") son presentados en secciones, organizadas según los objetivos de la evaluación. Al final, se sistematizan las sugerencias de perfeccionamiento para el programa ("Lo que puede ser hecho para mejorar el programa").

Los reportes de evaluación de la ejecución de los programas federales, conducidos por la Contraloría General de la Unión, en forma de sorteos sistemáticos de estados y municipios, son también fuentes de información muy interesantes para evaluaciones preliminares del proceso de implementación de programas públicos. Un ejemplo es el de Vieira (2008), que empleó los reportes de la CGU para identificar los problemas de gestión municipal del programa de Asistencia Farmacéutica en una muestra de 597 municipios. La autora constató que el 90,3% de las localidades presentó problemas en la gestión de recursos o servicios, como falta o deficiencia de control de stock de medicamentos, compras irregulares y condiciones inadecuadas de almacenamiento entre otros aspectos.

Los documentos, informes y reportes de órganos internacionales de Naciones Unidas – PNUD, Unicef, Unesco, Habitat, FAO, OIT, Cepal, etc. – así como Banco Mundial, OCDE y Banco Interamericano de Desarrollo son también importantes fuentes de información para la evaluación de Políticas, programas y acciones gubernamentales en escala nacional e internacional. Estos informes, usualmente anuales, traen datos estadísticos, análisis temáticos y experiencias de diferentes países. Se pueden hacer comparaciones de indicadores sociales, considerando los presupuestos sociales y la dinámica económica de cada país. Se puede conocer experiencias y programas de otros países, lo que son las mejores prácticas, lo que son los retos y dificultades comunes. Hay que considerar, como en toda y cualquier evaluación o estudio, que la visión de mundo y los enfoques metodológicos prevalecientes de los autores y de la institución ejecutora determinan los resultados que se hayan escrito en los reportes. Aun así, se pueden hacer estudios evaluativos útiles para proveer conocimiento interno a las oficinas y gestores de programas.

Cuadro 3: Algunos de los Principales reportes sociales internacionales

Institución	Informe Social
CEPAL	Panorama Social de América Latina y Caribe
PNUD	Informe de Desarrollo Humano
UNESCO	Global Education Digest
UNICEF	Estado Mundial de la Infancia
FAO	El estado mundial de la agricultura y la alimentación
OIT	Tendencias mundiales del empleo
Banco Mundial	Informe sobre Desarrollo Mundial
OCDE	OECD Factbook
Bco Interamericano	Development Effectiveness Overview

Consideraciones finales

La ampliación del gasto social en Brasil y otros gobiernos en Iberoamérica y la diversificación de los programas dirigidos a atender las diversas demandas públicas vienen presionando al sector público para refinar sus prácticas de gestión. En este sentido, está creciendo la preocupación con el perfeccionamiento técnico en la elaboración de diagnósticos y en las actividades de monitoreo y evaluación de programas⁶.

Estructurar mejor los sistemas de indicadores de monitoreo y especificar investigaciones de evaluación más consistentes son los desafíos que deben ser rápidamente enfrentados en los tres niveles de gobierno —federal, provincial y municipal—, bajo pena de extender por más tiempo la superación de las desigualdades sociales en el país y bajo el riesgo de perder la convicción de que los programas públicos pueden ser vectores impactantes en el cambio social en Brasil.

Este texto pretendió mostrar que dichos instrumentos necesitan ser especificados de acuerdo con el nivel de madurez del programa, respondiendo a las demandas de información del gestor en las diferentes fases del “ciclo de vida” del programa. No existe una receta única, lista y concluida para diseñar estos instrumentos. Existen experiencias, recomendaciones y buenas prácticas. ¡Que ellas nos sirvan para enfrentar los retos de la estructuración de nuestros sistemas de M&E!

Referencias bibliográficas

- ALMEIDA, A.C. *Como são feitas as pesquisas eleitorais e de opinião*. Rio de Janeiro, FGV, 2002.
- ATKINSON, T. *et ál. Social Indicators: the EU and Social Inclusion*. Oxford, Oxford Univ. Press, 2005.
- BABBIE, E. *Metodos de pesquisa de Survey*. Belo Horizonte: Ed.UFMG, 1999)
- BENNETT, C.A. ; LUMSDAINE, A. A. *Evaluation and experiment: some critical issues in assessing social programs*. New York: Academic Press, 1975.

BOWEN, C.C.; BOWEN, W.M. Content Analysis. In: YANG, K. ; MILLER, G.J. *Handbook of research methods in Public Administration*. New York, CRC Press, 2008, p.689-703.

CANO, I. *Avaliação de programas sociais*. Rio de Janeiro: FGV, 2002.

CAREY, G. W. Government. *Encyclopedia Americana*. Danbury: Scholastic, 2006, p. 127-131.

CARVALHO, S.N. *Avaliação de programas sociais: balanço das experiências. São Paulo em Perspectiva*, São Paulo, 17(3-4): 185-197, 2003.

COHEN, Ernesto; FRANCO, Rolando. *Avaliação de projetos sociais*. Petrópolis, RJ: Vozes, 1999.

COTTA, T. C. *Metodologias de avaliação de programas e projetos sociais: análise de resultados e impacto. Revista do Serviço Público*, Brasília, 49(2): 103-124, 1998. *E contribuição para o debate. São Paulo em Perspectiva*, São Paulo, 17(3-4): 185-197, 2003

COUTINHO, R. G. *Subsídios para organizar avaliações da Ação Governamental. Textos para discussão*, Brasília, n. 776: IPEA, 2001.

CUNILL GRAU, N.; OSPINA BOZZI, S. *Fortalecimiento de los sistemas de monitoreo y evaluación en América Latina*. Washington, Banco Mundial/CLAD, 2008.

FREY, Klaus. *Políticas Públicas: Um debate conceitual e reflexões referentes à prática da análise de políticas no Brasil*. Planejamento e Políticas Públicas, 21. Brasília: IPEA, 2002.

GENERAL ACCOUNTING OFFICE. *Social research and development of limited use to national policymakers: Report to the congress Washington*, 1977.

GUIMARÃES, J.R.S; JANNUZZI, P.M. IDH, *Indicadores sintéticos e suas aplicações em políticas públicas: uma análise crítica*. Revista Brasileira. Est. Urbanos e Regionais, Salvador 7 (1):73-89, 2005.

HOWLETT, M.; RAMESH, M. *Studying public policy: policy cycles and policy subsystems*. *DonnMills: Oxford University Press*, 2003.

JANN, W.; WEGRICH, K. *Theories of the policy cycle*. FISCHER, F. et ál. *Handbook of public policy analysis: theory, politics and methods*. Boca Raton: CRC Press, 2007, p. 43-62.

JANNUZZI, P. M. *Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no Brasil*. Revista do Serviço Público, Brasília 56 (2): 137-160. Abr./Jun. 2005.

JANNUZZI, P. M., MIRANDA, W. L Pradin y Monit: *ferramentas para tomada de decisão no ciclo de políticas públicas*. Salvador, Boletim de Estatísticas Públicas, n.4, 2008, p.62-73.

JANNUZZI, P. M., MIRANDA, W. L., SILVA, D. S. G. "Análise Multicritério e Tomada de Decisão em Políticas Públicas: Aspectos Metodológicos, Aplicativo Operacional e Aplicações". Belo Horizonte: Revista Informática Pública, ano II, 2009, pg. 69-87.

LAVILLE, C. ; DIONNE, J. *A construção do saber*. Porto Alegre: Artmed/ED. UFMG, 1999.

LINDBLOM, C. *Todavía tratando de salir del paso*. SARAVIA, E. ; FERRAREZI, E. (orgs.) *Políticas públicas: coletânea*. Brasília: ENAP, 2006, p.99-122.

MACKAY, K. *How to build M&E Systems to support better government*. Washington, Banco Mundial, 2007.

MADEIRA, F.R. *A improvisação na concepção de programas sociais: muitas convicções, poucas constatações – o caso do primeiro emprego*. São Paulo em Perspectiva, São Paulo, 18(2): 78- 94, 2004.

MCDONALD, J. HAWTHORN.L. *Program Evaluation and Performance Measurement: An Introduction to Practice*. Thousand Oaks; Sage, 2006.

MILLER, D.C. *Handbook of research design and social measurement*. Newbury Park: Sage, 1991.

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO. *Guia metodológico para construção de indicadores para o PPA*. Brasília, 2009.

MOHR, L. B. *Impact analysis for program evaluation*. Thousand Oaks: Sage, 1995.

NASCIMENTO, B. *A reforma universitária de 1968: origem, processo e resultados de uma política pública para o ensino superior*. Tesis de Doctorado, UFRJ, 1991.

OWEN, J.M. *Program evaluation: forms and approaches*. New York> Guilford Press, 2007.

PAES-SOUSA, R.; VAITSMAN, J. *Metodologias e instrumentos de pesquisas de avaliação de programas do MDS: Bolsa Família, Assistência Social, Segurança Alimentar e Nutricional*. Brasília, DF: MDS; SAGI, 2007.

PARADA, E.L. Política y políticas públicas. SARAIVA, E. ; FERRAREZI, E. (orgs.) *Políticas públicas: coletânea*. Brasília: ENAP, 2006, p.67-95.

REZENDE, L. M. e JANNUZZI, P. M. *Monitoramento e Avaliação do PDE: IDEB e painel de indicadores*. Revista do Serviço Público, Brasília, 59(2): 121-150, 2008.

RICHARDSON, R.J et ál. *Pesquisa social: métodos e técnicas*. São Paulo: Atlas, 2002.

ROSSI, P. H. et ál. *Evaluation: a systematic approach*. Thousand Oaks: Sage, 2004.

RUA, M. G. *Algumas considerações preliminares sobre política e políticas públicas*. In, RUA, Maria das Graças e CARVALHO, M.I.V. (orgs). *O estudo da política: tópicos selecionados*. Brasília: Paralelo 15, 1998.

RUA, M.G. *A avaliação no ciclo da gestão pública*. Apostila do curso de Especialização em políticas públicas da Educação com Ênfase em monitoramento e Avaliação – MPA. ENAP. MIMÉO, 2004.

SARAVIA, E. ; FERRAREZI, E. (orgs.) *Políticas públicas: coletânea*. Brasília: ENAP, 2006.

SARAVIA, E. *Introdução à teoria da política pública*. In: SARAIVA, E. ; FERRAREZI, E. (orgs.) *Políticas públicas: coletânea*. Brasília: ENAP, 2006, p. 21-42.

SCANDAR NETO, W. et ál. *Sistemas de indicadores ou indicadores sintéticos: do que precisam os gestores de programas sociais*. In: HENRIQUES, R. et ál. (Org). *Educação na diversidade: como indicar as diferenças?* (Coleção Educação para Todos, Série Avaliação). Brasília, MEC/Unesco, n. 8, v. 25, p. 16-34. 2006.

TRIBUNAL DE CONTAS DA UNIÃO. *Avaliação do TCU sobre o Programa Nacional Biblioteca na Escola*. Brasília: TCU, 2002.

VIEIRA, Fabiola Sulpino. *Qualificação dos serviços farmacêuticos no Brasil: aspectos inconclusos da agenda do Sistema Único de Saúde*. Revista Panamericana de Salud Publica, vol. 24, n. 2, p. 91-100, 2008.

VILLANUEVA, L.F.A. *Estudio introdutorio*. SARAIVA, E. ; FERRAREZI, E. (orgs.) *Políticas públicas: coletânea*. Brasília: ENAP, 2006, p.43-65.

WEISS, C. *Evaluation research*. New Jersey: Prentice Hall, 1972.

WORTHERN, B. R. et ál. *Avaliação de programas: concepções e práticas*. São Paulo, EdUsp/Gente, 2004.

ⁱ Ver, por ejemplo, las publicaciones, eventos y sitios de las asociaciones profesionales o académicas de evaluación en estos países, entre los cuales: American Evaluation Association (www.eval.org); Canadian Evaluation Society (www.evaluationcanada.ca); European Evaluation Society (www.europeanevaluation.org); y Société Française d'Évaluation (www.afe.asso.fr).

ⁱⁱ De hecho, los trabajos de evaluación de programas públicos y sus autores se encuentran dispersos por las principales asociaciones científicas disciplinarias de Brasil, como: ANPAD (administración y administración pública); ANPEC (economía); ANPOCS (ciencias sociales); ANPED (educación); ABEP (estudios poblacionales); ABET (estudios del trabajo); ABRASCO (salud colectiva). Vale destacar algunos de los principales periódicos donde se pueden encontrar estudios evaluativos: Revista del Servicio Público; Revista Brasileña de Ciencias Sociales; Revista de Administración Pública; Revista Sao Paulo en Perspectiva; Textos de Discusión (IPEA); Planificación y Políticas Públicas; Investigación de Planificación Económica; Texto de Discusión ENCE, entre otros (algunos periódicos están disponibles en el portal www.scielo.br). A partir del año 2000, fueron creados espacios más multidisciplinarios para discusión y presentación de estudios de evaluación de políticas públicas, entre ellos la Asociación Brasileña de Evaluación Educativa (www.abave.org.br); la Red Brasileña de Monitoreo y Evaluación (<http://redebrasileirademea.ning.com>); el portal Políticas Públicas en Foco (www.boletimfundap.cebrap.org.br), mantenido por la FUNDAP y CEBRAP; el sitio del Tribunal de Cuentas de la Unión (www.tcu.gov.br) y de la Contraloría General de la Unión (www.cgu.gov.br), además de los foros creados por los programas de posgrado *lato y stricto sensu*, académicos y profesionales, y de las líneas de investigación en este sentido en programas de posgrado en administración pública, economía social, salud colectiva, derecho, etc.

ⁱⁱⁱ Como presentado en varios textos reunidos en la valiosa recopilación organizada por Saravia y Ferrarezi (2006) en políticas públicas, publicada por la ENAP y disponible en su sitio (www.enap.gov.br), el Ciclo de Políticas Públicas puede ser descrito con un número mayor o menor de etapas. De manera general se distinguen por lo menos tres macroetapas: formulación (en la que estaría la formación de la agenda), implementación y evaluación. Esta publicación organizada según las etapas del ciclo, presenta artículos clásicos que profundizan la vasta discusión en cada una de ellas.

^{iv} La resistencia al cambio puede ser incluso legítima en el sector público, cuando se intentan proyectos poco claros. Además, la incorporación de innovaciones no es problema sólo en el sector público, sino también en las empresas privadas. Los mecanismos de motivación, presión y coerción en estas últimas acaban movilizandolos cambios, a veces a costos personales e institucionales elevados.

^v Jannuzzi (2005), disponible para descargar en la Revista del Servicio Público, v.56, n.2, en www.enap.gov.br.

^{vi} Esta publicación del IBGE, así como otras de la misma institución, están disponibles para descargar en www.ibge.gov.br.

^{vii} Una discusión de los límites y potencialidades del IDH puede ser encontrada en Guimarães y Jannuzzi (2005).

viii Un buen ejemplo de indicadores para diagnósticos de programas sectoriales en Brasil es demostrado en la publicación “Construyendo el diagnóstico municipal”, disponible en www.cepam.sp.gov.br. Otro ejemplo de diagnóstico apoyado en indicadores multitemáticos es el Diagnóstico para Programa de Calificación Profesional del Estado de Sao Paulo, disponible en www.emplo.sp.gov.br.

ix En el sitio www.anipes.org.br está disponible un aplicativo (Pradin) que permite entender el potencial del Análisis Multicriterio en la toma de decisiones en políticas públicas.

x Las investigaciones y datos de estas instituciones pueden ser consultadas, respectivamente, en www.ibge.gov.br, www.datasus.gov.br, www.inep.gov.br. En el portal www.anipes.org.br, pueden ser accedidos sitios de órganos subnacionales de estadística en Brasil.

xi Ver lo que las agencias estadísticas americanas producen regularmente en este sentido, accediendo el portal www.fedstats.gov. El Bureau of Labor Statistics produce, por ejemplo, estimativas mensuales de tasa de desempleo para 372 áreas metropolitanas y tasas anuales para los más de 3 mil condados y 50 mayores ciudades americanas. El Bureau of Census, por su parte, presenta anualmente estimativas poblacionales para más de 20 mil localidades e indicadores sociales referidos a más de 10 temas —desde ocupación, vivienda a educación— para municipios por arriba de 65 mil habitantes.

xii Ver los aplicativos de la Senarc/MDS (www.mds.gov.br) en este sentido, los sistemas de consulta *online* de la RAIS y del CAGED en el MTE (www.mte.gov.br), y los datos de la Previsión Social en la herramienta Infologo en www.dataprev.gov.br.

xiii En el referido texto, se empleó el aplicativo Monit —disponible en www.anipes.org.br— para componer un panel situacional con gráficos de los indicadores seleccionados para dos unidades territoriales de interés de monitoreo. Una breve presentación del aplicativo es encontrada en Jannuzi y Miranda (2008), disponible en el Boletín de Estadísticas Públicas n.4, en el mismo sitio.

xiv Consulte el “Diccionario de variables, indicadores y programas sociales” en <http://aplicacoes.mds.gov.br/sagi/dicivip>. Este diccionario también está disponible para descargar en www.mds.gov.br/sagi.

xv Acceda estos portales, respectivamente, por medio de un buscador en la Web, buscando Siga Brasil + Senado; dirección <http://aplicacoes.mds.gov.br/sagi>; sitio www.emplo.sp.gov.br; sitio www.agenda2012.com.br.

xvi Un paquete estadístico gratuito de amplia difusión es el Epi-info, disponible en www.lampada.uerj.br, donde se puede encontrar también material de consulta y entrenamiento. El Weka, disponible en www.cs.waikato.ac.nz/ml/weka, es un aplicativo gratuito para extracción de datos, es decir, dispone de rutinas para análisis descriptivo y exploratorio de datos.

xvii Los sitios citados poder ser accedidos por www.mds.gov.br/sagi, publicaciones; www.nadd.prp.usp.br/cis.

^{xviii} Los reportes de esas instituciones pueden ser accedidos en www.tcu.gov.br, menú "Evaluación de Programas de Gobierno" y www.cgu.gov.br, menú "Auditoría y Fiscalización".