



UNIVERSIDAD DE BELGRANO

Documentos de Trabajo

Area de Estudios de la Educación Superior

**La implementación de estudios de evaluación
retrospectiva: problemas y perspectivas en la
Educación Superior Pública**

Nº 112

Héctor R. Gertel

Departamento de Investigación

Julio 2003

Una innovación institucional mayor, recientemente introducida en muchos países es la aplicación de indicadores de resultados en la educación superior pública, presentada como la respuesta más apropiada ante reclamos crecientes de la sociedad civil por mayor publicidad de los actos de gobierno en general y de los agentes educativos, más específicamente. El trabajo pone a prueba esta visión popular y se pregunta: ¿Quién se beneficia de esta innovación institucional? Utilizando un enfoque del tipo principal-agente, concluye que el desarrollo de este tipo de innovaciones responde al interés del principal, en este caso la autoridad educativa, preocupada por generar información para su propio dominio. De manera que la presunción original de una fuerte asociación de la innovación con reclamos de los agentes que conforman la sociedad civil parece debilitarse. La sección primera introduce el marco de análisis. Luego, se señalan los puntos centrales al desarrollo de indicadores de resultado. El tratamiento de indicadores de resultado en la educación superior se ilustra a partir de la potencialidad que crea en la Argentina el establecimiento de la Comisión Nacional de Evaluación y Acreditación Universitaria en 1995. La sección final concluye con una breve discusión de los resultados, que confirman el sesgo que las técnicas de evaluación utilizadas introducen a favor de las necesidades del principal.

1. Propósitos y alcances del trabajo

Los estudios de evaluación retrospectiva han alcanzado un alto grado de aprobación y difusión dentro del ámbito de las decisiones públicas. Se trata de un tipo de ejercicio derivado de la literatura sobre indicadores de resultados con elementos de la teoría de la elección pública. Los ejercicios de evaluación retrospectiva han alcanzado un alto grado de aprobación y difusión dentro del ámbito de las decisiones públicas. El presente trabajo se concentra en los problemas y los potenciales beneficios de implementar estudios de evaluación retrospectiva en la educación superior pública, un ámbito siempre difícil en la relación principal-agente y busca obtener algunas lecciones de ello.

La introducción de los indicadores de resultados en el campo de la educación superior pública es un fenómeno relativamente nuevo que, en parte, responde a los crecientes reclamos de la sociedad por una mayor publicidad de los resultados derivados de la aplicación de los dineros públicos. Evaluar el desempeño de los agentes ha sido tradicionalmente considerado una tarea que entraña profundas complicaciones dentro de la educación superior. En la educación superior pública, la implementación de indicadores de resultados reconoce además la preocupación por medir el impacto de las innovaciones institucionales impulsadas por parte de la autoridad educativa en casi todos los países durante el último cuarto de siglo. Reformas que, principalmente, están destinadas a mejorar la eficiencia y la calidad sistémica de la oferta, promoviendo, incluso, su diversificación, de manera de atender así mismo objetivos de equidad. Los orígenes de este proceso se remontan a los años setenta, cuando numerosos países dentro de la Comunidad Europea comenzaron a experimentar fuertes presiones por expandir y diversificar las opciones en la educación superior (OECD, 1999). También contribuyó a ello el interés de las agencias internacionales por medir retrospectivamente el impacto de sus operaciones de préstamos para el desarrollo institucional de la educación superior en países menos desarrollados, inaugurados en el decenio de los sesenta (Gertel, 1986; Gertel, 1984). En la Argentina, los programas de innovación institucional son de más reciente data. Entre los mismos sobresalen la creación de la Comisión Nacional de Evaluación y Acreditación Universitaria (CO-NEAU) en 1995 y, a partir de 1997, las pautas de distribución presupuestaria para las Universidades Nacionales basada en fórmulas (Delfino y Gertel, 1996). En ambos casos, los indicadores de resultados están llamados a jugar un papel principal dentro de la evaluación retrospectiva ya que estos nuevos procedimientos permitirían, al menos en parte, superar una etapa de desarrollo universitario en la cual, tanto en la Argentina como en otros numerosos países los intereses particulares propios de las universidades públicas llegaron a primar por sobre el interés social (Williams, 1994:pp3-7). De ahí que, siempre dentro de este enfoque particular, por comportarse de este modo, los principios tradicionales aplicados a los agentes privados, no podría aplicarse a los agentes públicos. La razón de ser de la teoría de la elección pública es, precisamente, explicar de que manera este supuesto erróneo puede corregirse y que implicancias se derivan para las políticas públicas. La conducción de estudios de evaluación retrospectiva surge en respuesta a la necesidad de conocer mejor la conducta de los agentes públicos y su impacto en términos de resultados auditables y eficacia en la gestión de los recursos.

Los indicadores de resultados constituyen una herramienta central de la evaluación retrospectiva. Para la universidad pública, el interés creciente en los indicadores de resultados tiene como trasfondo los fuertes cambios en la política de la educación superior en la Argentina y otros países. En la Argentina, durante los últimos 15 años, una serie de estudios originados en el Ministerio de Educación y otras áreas oficiales y privadas y la aparición de la Ley Federal de Educación y la Ley de Educación Superior alteraron la manera tradicional de hacer política de la educación superior en el país, incluida la manera de orientar las pautas presupuestarias hacia las universidades nacionales y hacia los colegios de nivel terciario. Nuestros socios en el Mercosur, también Chile, pero principalmente los países miembros de la OECD confrontan una realidad similar, con acrecentadas demandas del público por conocer más acerca de la educación superior pública, cuanto y en que gasta, y con que resultados. De ahí el creciente interés que despiertan los estudios de evaluación retrospectiva en la esfera de las decisiones públicas.

2. Puntos centrales en el desarrollo de indicadores de resultado

Luego de transcurrido un quinquenio desde la controvertida introducción de los indicadores de resultados en la política universitaria argentina, su continuidad se encuentra en un punto crítico. Desde su surgimiento como una iniciativa de política, en gran medida externamente impuesta a las universidades, los indicadores de logros han sido gradualmente aceptados e incorporados a varios niveles dentro del gobierno y administración de las instituciones universitarias, por ejemplo, mediante la introducción de la auto-evaluación institucional, voluntaria en la mayor parte de las experiencias realizadas hasta el presente. Otro ejemplo lo constituye el debate sobre los indicadores que deberían orientar el Programa de incentivos a docentes-investigadores. Luego de una etapa inicial extenuante de divulgación, los indicadores inicialmente introducidos en los mencionados programas resultan hoy mejor comprendidos, por un mayor número de agentes involucrados, y su desarrollo, aún incipiente, ha demandado del sistema la aplicación de fuertes sumas de recursos. Al mismo tiempo, la incertidumbre, tanto en relación a comparabilidad, como en cuanto a la utilización potencial y significado último de este ejercicio de evaluación retrospectiva se han intensificado.

El desarrollo de los indicadores de resultado como instrumento de política pública es, cada vez más, considerado como el producto combinado de los avances técnicos en la materia y los intereses políticos de la agenda de las propias universidades públicas. El marco dentro del cual se analicen las diferentes aplicaciones de los indicadores de resultados en la educación superior debe contemplar el reconocimiento de esta realidad. (Kogan, 1993, Cave y Hanney, 1992). Dentro de este marco, una utilización diferenciada de un mismo indicador bajo circunstancias diversas debería poder explicarse en términos de cuatro variables: (i) desarrollo técnico, (ii) decisión política de crear y sostener adecuadamente las estructuras que permitan e incentiven su uso, (iii) decisiones políticas de incorporar cambios en función de las señales que proveen los indicadores, y (iv) el balance de influencias que es capaz de ejercer cada una de las partes involucradas sobre la redacción de las recomendaciones y su implementación.

Técnicas cada vez más avanzadas están siendo aplicadas en la Argentina y en otros países para medir e interpretar los impactos de las políticas en la educación superior y en los sectores sociales en general, discriminando por programas (OECD, 1999, Jimenez y Sawada, 1999); aunque con incierto impacto sobre la revisión de las decisiones, se trata de una etapa superadora de otra, anterior, donde la automaticidad en los incrementos (o reducciones) presupuestarios era el fenómeno más significativo. Por un lado, hay en la actualidad un proceso en marcha por el cual se espera que los indicadores resulten a la vez más robustos y específicos al programa que se desea resaltar. De todas maneras, el mismo proceso expuso lo complejo que resulta avanzar en esta dirección. En parte, aunque no totalmente por estas razones, se acrecientan los debates acerca del lugar que debe asignarse a los procesos de evaluación retrospectiva en base a indicadores de resultados dentro de la política pública sobre la educación superior.

No-solo el contexto técnico, también el político está en proceso de cambio. Los gobiernos intentan respuestas a las crecientes y muchas veces contrapuestas demandas que provoca sobre la educación superior la transformación acelerada de la economía adhiriendo a principios de gerenciamiento de los servicios públicos afines a los del sector privado en la forma de, por ejemplo, incorporación de conceptos de calidad total en el manejo del gobierno u prestación de servicios en base a señales de la demanda (vales o

subastas). En el contexto de los actuales rápidos cambios económicos se asume que el conocimiento pasó a ocupar un lugar central y creciente en la formación del capital de una sociedad (OECD, 1989: 56-65). Probablemente por este hecho, los sistemas de educación superior están creciendo nuevamente en el país, luego de una década con relativa estabilidad desde 1987 a 1996. En este proceso se está acrecentando la diferenciación entre opciones que ofrece la educación superior, las partes involucradas son más y más variadas, la población estudiantil ha diversificado sus intereses y otros grupos de la sociedad desean consumir nuevos tipos de servicios, sea de capacitación asociada con nuevos empleos o con la investigación.

Si la fuerza inicial para incorporar indicadores de eficiencia en la educación superior en países como la Argentina provino de visiones de gobierno comprometidas con la mejor utilización del subsidio público a las universidades y para establecer un vínculo más estrecho y de mayor transparencia pública entre financiamiento y resultados en el marco de una angustiante estrechez de recursos, gradualmente otros valores además de costos y eficiencia han entrado a terciar. Hoy, mucha gente dentro de la educación superior puede percibir las ventajas de conocer los procedimientos, tablas de valoración, incluso nombres de evaluadores, con los que se ha determinado el valor relativo de programas, proyectos y actividades que los involucran personalmente, y también otras voces interesadas acceden a la misma información y forman su propia composición de lugar. Pero diferentes perspectivas, podrían determinar una multiplicación de apreciaciones en relación a las prioridades (Levin, 1974, Rivlin, 1971:9-45).

Por fin, la creciente aceptación que han alcanzado los indicadores de resultados es parte de un cambio actitudinal en favor de un ambiente de trabajo más competitivo que ha comenzado a manifestarse en las universidades públicas de la Comunidad Europea durante la última década. Pero está claramente establecido que habría pocos indicadores de aplicación universal por lo que su utilización mas bien aparece como parte de un movimiento para asegurar una mayor calidad y control de los procesos a partir de ciertos resultados reconocidos y validados (Kells, 1993). En la Argentina, este mismo fenómeno ha comenzado a manifestarse en la relación entre el gobierno nacional y las universidades públicas, por ejemplo, en relación a las carreras de riesgo (Morera y Brissón, 1998) mientras que las jurisdicciones provinciales registran una deuda con respecto a la evaluación retrospectiva de los colegios terciarios.

3. Indicadores de resultados de la educación superior

La universidad pública en particular y la educación superior en general corresponde al modelo de una firma multi-producto que básicamente produce, bajo modalidades diferenciadas, enseñanza e investigación en varias plantas de producción cuya localización influye sobre los resultados esperados. En este contexto el indicador de resultados es un valor numérico que busca medir algo que es difícil de cuantificar: el volumen de docencia impartida de una determinada calidad, o los frutos de los programas de investigación. Por ejemplo, un ratio entre insumos y producto satisface esta definición, pero un índice de satisfacción que ordena aspectos cualitativos desde la perspectiva de los agentes no debería ser desdeñado a-priori. En suma, los indicadores de resultados de la educación no necesariamente responden a un ordenamiento cardinal estricto en un sentido matemático, pero se espera que provean información para discriminar una mejor opción de otra (Levin, 1983: 11). De hecho, la diferencia entre un indicador general y un indicador de resultados estriba en que estos últimos son un punto de referencia, por ejemplo, un standard, un objetivo, un índice de comparación y son de carácter relativo más que absoluto lo que entraña adosarle un juicio de valor. Hay otro tipo de indicadores de resultados que vinculan información externa con una apreciación de éxito en el trabajo interno de la institución (por ejemplo, el trabajo del Fondo de Mejoramiento de la Calidad -FOMECA, es anualmente evaluado por un Comité Internacional de Seguimiento que hace público su Informe).

Los indicadores de resultados pueden ordenarse de la manera convencional según una secuencia del tipo insumo-producto, como la desarrollada en Levin (1983). La evaluación retrospectiva según este autor implica cubrir todas las partes de un proceso de decisión previo: Fue el problema bien identificado?, ii. Se consideraron todas las alternativas?, iii. Listado de los ingredientes, iv. Valuación de los mismos, v. Asignación de costos a diferentes partes interesadas, vi. Distribución de beneficios y efectos, vii. Ratio de beneficios, o efectividad a costos. Así, la evaluación retrospectiva en la universidad puede verse como un proceso que analiza el tiempo de los estudiantes, el tiempo de los profesores, los demás insumos y los servicios de

equipos y edificios, y con que eficacia éstos son transformados en productos relacionados con la docencia y la investigación (por ejemplo, en Delfino y Gertel [1996] se presenta un ejercicio de evaluación retrospectiva sencillo donde se ilustra el uso de costos estándar de docencia para la decisión).

Hay un método alternativo para comparar insumos y productos entre instituciones, conocido como la «frontera de eficiencia», que marca el mejor índice de producción que es posible alcanzar utilizando la mejor práctica tecnológica conocida. En educación, los límites de eficiencia resultan bastante difusos puesto que difícilmente puedan incluirse acabadamente todos los ingredientes intervinientes. Pero es un enfoque utilizado para comparar la efectividad relativa de instituciones similares y establecer cual de ellas sobrepasa en efectividad a las demás y porqué (Levin, 1997). En Delfino y Gertel [1996] se provee una ilustración sencilla de las posibilidades de este enfoque dentro de una evaluación retrospectiva de las universidades públicas en la Argentina.

Las características de «autoselección» de la población estudiantil a causa de la deserción afecta los resultados de la enseñanza en la universidad pública y en la educación en general. Recientemente, la evaluación retrospectiva ha comenzado a prestar una mayor atención a los problemas de sesgo en la medición que introduce el fenómeno de la autoselección. Sea como somera ilustración el caso de un individuo que selecciona el nivel de estudios que brinda el mayor retorno, V_j . Hay dos opciones: $j=U$ si está matriculado en la universidad con un programa nuevo de becas- arancel, $j=T$ si no está matriculado en esa universidad. Bajo que circunstancias distintos individuos seleccionarían la carrera con un programa nuevo? Sea U_j que depende de los beneficios y costos relativos de los estudios universitarios que serán seleccionada si

$$U_i^* = V_{ui} - V_{Ti} > 0$$

donde U^* es una variable latente que describe la probabilidad que una persona i seleccione proseguir estudios en la universidad con el nuevo programa. Pero como una misma persona i no puede encontrarse al mismo tiempo en el grupo de inscriptos y de no inscriptos, es necesario recurrir a estimaciones indirectas de cambios en el tiempo para grupos de individuos con idénticas características. Pero los cambios en el tiempo pueden deberse no solo a variaciones entre el grupo bajo estudio y el grupo control en el momento de la evaluación, sino también a variaciones en ambos grupos a través del tiempo. Así, la evaluación retrospectiva debe controlar ambos efectos separadamente

$$U_i^* = [V_{ui} - V_{Ti}]_{T=t} - [V_{ui} - V_{Ti}]_{T=0}$$

Luego, se estima independientemente $[V_{ui}]_{T=t}$; $[V_{Ti}]_{T=t}$; $[V_{ui}]_{T=0}$; $[V_{Ti}]_{T=0}$ utilizando técnicas probit de la manera usual. Algunos ejercicios importantes que rescatan estos aspectos han sido desarrollados en Jimenez y Zawada, (1999) y en Kim, Alderman y Orazem (1999). La identificación del efecto neto de «autoselección» en la universidad pública confronta dificultades serias porque los sistemas estadísticos no están preparados para captar la información requerida.

4. La preparación de los informes de evaluación retrospectiva: Propósitos y Alcances.

La preparación de los informes de evaluación retrospectiva debe rescatar que el principal propósito de la evaluación retrospectiva es apoyar el desarrollo técnico y mejorar el análisis del impacto del programa. La bondad y pertinencia de la información disponible juegan un papel central en ello. De aquí que el informe tratará de sintetizar el voluminoso trabajo que surge de aplicar, por ejemplo el esquema de trabajo de «insumo-producto» y responderá a preguntas tales como:

- (a) ¿Se seleccionó una forma de análisis socioeconómico más apropiada que otra?
- (b) ¿Cuan confiables son los resultados ex-ante del análisis socioeconómico de impacto?
- (c) ¿Qué papel se asignó a los estudios ex-ante en la toma de decisiones durante la ejecución del programa, y como esto afecta a los resultados esperados?
- (d) ¿Es conveniente recomendar trabajos para aumentar la confiabilidad de los indicadores?

Sin acceso a la información que los estudios retrospectivos proveen, hay poca base práctica para el desarrollo de lineamientos para conducir nuevos estudios de impacto, o para efectuar recomendaciones orientadas a mejorar los procesos del trabajo diario que afectan a los resultados. Por lo tanto, los estudios retrospectivos contribuyen a mejorar el valor de los análisis socioeconómicos ex-ante para los decisores públicos.

Los alcances del análisis retrospectivos se extienden:

- (a) a la provisión de información para realimentar la realización de nuevos estudios socioeconómicos, ayudando a mejorar su consistencia;
- (b) a las oportunidades para aprender de experiencias del pasado y aplicar las lecciones a análisis futuros;
- (c) a la identificación de áreas que requieren mayor investigación o desarrollo, por ejemplo, cuando el estudio pone de manifiesto tipos y categorías de impactos que han sido previstos con un menor grado de acierto o precisión, y
- (d) estimulando mejoras y refinamientos en las técnicas de análisis.

Son escasos los ejemplos de estudios retrospectivos que cubren satisfactoriamente todos los aspectos indicados. Entre las dificultades coincidentes se destacan:

- (a) Dificultades para cuantificar explícitamente todos los impactos. Esta limitación puede derivar muy probablemente en el uso ineficiente de los recursos nacionales. Este aspecto ha sido estudiado en Levin (1997) en Estados Unidos, para la educación primaria, y en Hanushek (1994) para la educación pública de ese país en general. Se ha sugerido, al menos para países de mediano nivel de desarrollo, que una cobertura alta de la educación superior en relación al grupo de edad (¿quizá por encima del 40 por ciento?) pudiera conducir a costos de congestión que afectarían negativamente la contribución de las universidades al crecimiento económico. Pero los resultados obtenidos no son muy conclusivos.
- (b) Diferencias importantes entre los costos efectivos y los previstos. Algunas de las causas pueden estar asociadas con una incorrecta predicción en la evolución de la población estudiantil, o del empleo y los salarios, causada por distorsiones en la información que se facilita al evaluador, la que fue construida sobre la base de apuestas estratégicas de diferentes actores interesados, sea en pronosticar un aumento más acelerado de lo real, o en ocultarlo.
- (c) La predicción, cuantificación y valuación de los beneficios puede aún resultar más problemática, porque, típicamente, la información necesaria para el control experimental de respuestas a tratamientos, la medición de tasas de éxito y otras relaciones, como "población atendida a población potencial total" es endeble o inadecuada, o, por fin, inexistente. Y pudiera no ser posible reconstruirla durante el estudio retrospectivo. Un segundo punto importante en esta aspecto es la debilidad de los términos de referencia, poco precisos al momento de solicitar el examen del nivel de incertidumbre que acompaña a los estimadores de beneficio.
- (d) Aun excluyendo los aspectos de formulación imprecisa y credibilidad de los indicadores de impacto, quienes los utilizan no están exentos de la tentación de sobreabundar en el detalle de los aspectos positivos y disimular los negativos en razón del temor a las repercusiones políticas y otras limitaciones del entorno; y
- (e) La resistencia a asociar/consultar a todos los participantes desde un principio puede más tarde causar problemas significativos a quienes toman decisiones y determinar desvíos importantes en el logro de los beneficios esperados.

Los estudios retrospectivos, asimismo, destacan que asociar a las partes involucradas, por ejemplo, a partir de un ejercicio de "autoevaluación" tiene un gran valor. Los mecanismos participativos pueden contribuir a incrementar los beneficios o reducir el costo del riesgo de una decisión desacertada o resistida de manera significativa, porque desde la posición de los excluidos, siempre se escuchará que sus justos reclamos no son atendidos, o que lo que se imputa como "logro" desde otras perspectivas, son costos para este grupo.

Desde luego que no resulta posible desarrollar estimadores monetarios de los beneficios o de la reducción de costos asociados con la realización de estudios de evaluación retrospectiva, pese a ello, algunos autores arriesgan decir que los beneficios podrían exceder hasta un tercio el costo de los estudios (Morgenstern, 1997). Pese al beneficio que presuntamente se desprende de los estudios de evaluación retrospectiva, su difusión está limitada por la dificultad de lograr fondos presupuestarios para los mismos y otros

problemas, entre los que se destaca la dificultad de recolectar, ex-post, la información que no se previó ex-ante y que necesita para estimar el real impacto que se desprende de un programa o acción, principalmente si éstos se conducen por objetivos, o en base a estándares. En razón de su alto costo, en general, se trabaja en base a muestras seleccionadas.

Tabla 1
Los aspectos a verificar en una propuesta de evaluación retrospectiva

A. Aspectos relativos a la fortaleza del Documento de Proyecto original

1. ¿Fue el programa considerado favorablemente por todos los interesados? (por ejemplo, ¿hubo controversias públicas y desacuerdos?)
2. ¿La decisión de llevar adelante el Programa se justificó en los estudios de base?, o ¿la misma estuvo influenciada por otros factores?
3. ¿Se logró un grado de consenso alto? (¿hasta qué punto las opiniones de las partes involucradas resultaban incompatibles o no?)
4. ¿Fue amplia la participación de los interesados?
5. ¿Se incluyó la evaluación por pares?

B. Aspectos relativos al estudio retrospectivo, propiamente

1. ¿Existe información apropiada y detallada para realizar el estudio?
2. ¿Están los interesados dispuestos a cooperar?
3. ¿Con qué urgencia deben ser elevados los resultados?
4. ¿Se encuentran disponibles los fondos que requiere el estudio?

La Tabla 1, resume la anterior discusión indicando el conjunto de factores que deberán ser considerados al preguntarse si un a propuesta de estudio retrospectivo arrojaría los beneficios esperados.

Los distintos tipos de estudios retrospectivos

La figura 1 provee una síntesis que facilita la comprensión de los distintos tipos de estudios retrospectivos. En función de los objetivos que éstos persiguen, es conveniente señalar tres tipos de estudios retrospectivos:

Estudios de seguimiento de la ejecución, indagan acerca de la calidad de los estudios de base, en particular la rigurosidad con que los mismos se realizaron, la calidad de la información analizada, el grado de precisión en las previsiones y una apreciación acerca de si las mismas podrían haber sido mejoradas desde la conducción durante la ejecución del Programa.

Estudios de impacto, donde el énfasis está puesto sobre los procesos de management, tratando de comprender si los estudios de base fueron utilizados para la decisión y en qué medida influyeron sobre los resultados de las acciones del Programa.

Estudios sobre la precisión de las proyecciones en términos de mínimas diferencias entre lo proyectado y lo realizado, tratando de descomponer la misma, individualizando los factores que determinan la diferencias y proponiendo mejoras en la calidad de los indicadores.

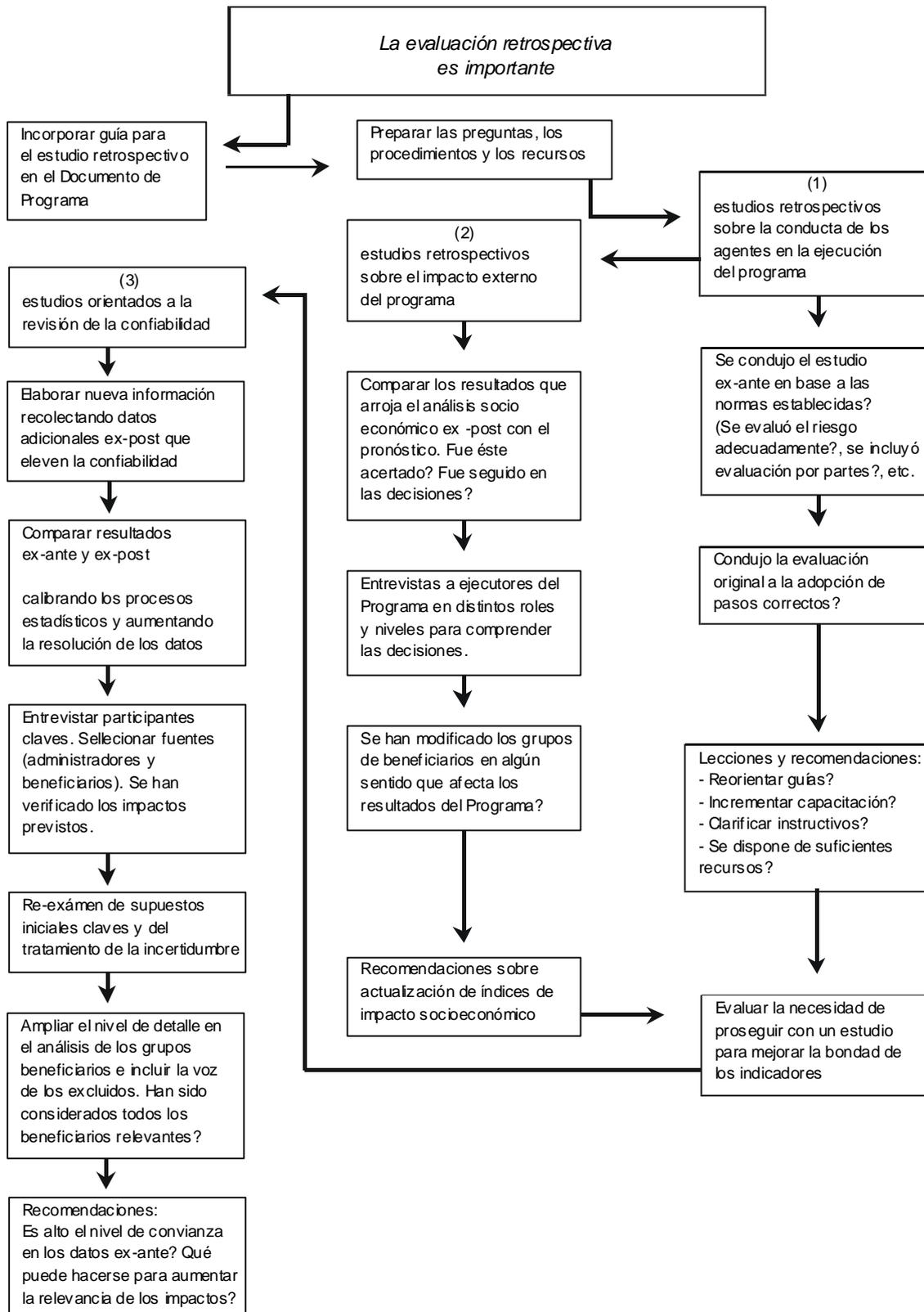
En la prácticas, los tres tipos de estudios están interrelacionados y los problemas básicos de la decisión se concentran en seleccionar el tipo de estudios a encarar.

Idealmente, es deseable contar con pautas indicativas para el estudio retrospectivo elaborada durante la preparación del Proyecto. Pero la situación más frecuente es que éstas pautas, si existen, no se hayan tomado en cuenta durante el desarrollo del Programa, ni en el plan de cuentas contables, ni en los gastos de operación, por lo que, típicamente, los datos con los que se contará al momento del estudio retrospectivo, o no existen, o fueron recolectados y agregados con otros propósitos, perdiéndose gran parte de su utilidad.

Desde luego, los costos de un estudio retrospectivo deben resultar justificados por los beneficios, sea en términos de una mejora en la conducta de los agentes o una mayor precisión en la identificación de los impactos socioeconómicos. Por fin, anticipar el potencial de los estudios retrospectivos en el diseño del Programa es, en sí mismo un elemento que ayuda a reducir costos, porque quienes conducen los estudios ex ante resguardarán información valiosa sobre fuentes de datos, procedimientos y decisiones que facilitan los estudios retrospectivos.

Una vez tomada la decisión de realizar un estudio retrospectivo, el análisis de la conducta de los agentes es usualmente el primero en ser desarrollado, pues ayuda a identificar las razones del grado de exactitud o de inexactitud que rodea a la información existente. Pero la importancia relativa de la inexactitud solo puede ser evaluada al término de los estudios de la conducta de los agentes o del impacto del programa. Debe, sin embargo, señalarse que muchas veces, los errores cometidos en la etapa del diseño y durante la implementación resultan tan notorios que la evaluación retrospectiva, previo a todo otro tipo de consideración, debe proceder a la revisión de la información disponible de modo de aumentar tanto su confiabilidad como los niveles de precisión.

Figura 1
Guía indicativa para procesos de evaluación retrospectiva



La importancia de una guía pre-proyecto para realizar evaluaciones retrospectivas

El contar con una buena guía para la realización de la evaluación retrospectiva no sólo contribuye a hacer ésta más sencilla, también recorta gastos innecesarios asociados con la improvisación. La guía usualmente incluye los aspectos siguientes:

(i)Cuál es el propósito del estudio retrospectivo; (ii) Cuánto tiempo después del estudio ex-ante se llevará a cabo el estudio retrospectivo; (iii) si se plantea realizar un solo estudio, quizá de gran magnitud, o se espera construir un proceso de evaluación continua para destacar tendencias o dificultades recurrentes; (iv) el rol y las responsabilidades de los participantes (por ejemplo, de quienes toman las decisiones en el ente regulador, de los agentes que las implementan, de los beneficiarios) que incluyen la distribución de los costos y las responsabilidades de ejecución y control; (v) gerenciamiento, que incluye decisiones sobre estrategias de participación amplia o restringida de todos los interesados, y (vi) la forma que tomará el Informe de Evaluación (a quién va dirigido, es público o reservado, etc.) es quizá, el aspecto más delicado. Puesto que en el documento queda expresado todo lo hallado, su publicación genera indudables compromisos, de ahí que el número de informes completados pero no publicados es demasiado grande restándole a estos estudios el potencial beneficio que de ellos se espera. Es necesario, entonces, dejar establecidas normas para su difusión.

Los estudios de evaluación retrospectiva sobre las conductas y los impactos

Tanto los estudios sobre la forma en que se condujo el Programa, como sobre sus impactos respetan los lineamientos generales arriba esbozados. La primera etapa en los estudios sobre la conducta de los agentes o sobre el impacto es determinar los objetivos y el alcance del trabajo. En los estudios sobre la conducta de los agentes, el principal objetivo es determinar si la preparación del programa fue correctamente orientada, y cuando se encontrara que no lo fue, indagar el porqué. En los análisis retrospectivos de impacto, el objetivo es identificar si el estudio socioeconómico previo se utilizó para tomar las decisiones y, si no, qué causas lo impidieron.

El alcance de un estudio sobre el comportamiento de los agentes varía dentro de muy amplios límites. Por ejemplo, en algunos casos sólo se solicita verificar si los agentes actuaron de acuerdo a las reglas establecidas. En otros casos, el estudio pudiera examinar los componentes del proceso de decisión (si se evaluaron correctamente los riesgos, si existió evaluación por pares, si hay mecanismos regulares de publicidad, etc.). En cada caso, los diferentes niveles de análisis pueden ser encargados a instancias previamente establecidas, tales como, departamentos de auditoría internos, oficinas de auditoría gubernamental o independiente, o expertos contratados. Desde luego, el mayor valor de estos estudios surge cuando es posible efectuar comparaciones entre los resultados de muchos de ellos de manera de poder observar regularidades comunes, bajo condiciones controladas ("meta-análisis). De hecho, revisar un diagnóstico socioeconómico en forma aislada difícilmente permita identificar si las divergencias entre los términos de referencia y lo realizado se debe a desvíos en la conducta del agente o se debe a una formulación deficiente de los términos de referencia, que quizá deban ser modificados.

Los estudios de conducta e impacto pueden descubrir que el análisis ex-ante, en algunas de sus partes se encuentra claramente errado o que los ejecutores adoptaron decisiones francamente cuestionables. Posiblemente, esto explica la escasa voluntad de autoridades políticas y técnicos para apoyar los estudios de evaluación retrospectiva en su área de decisión. Por ello, es necesario insistir que estos estudios apuntar a mejorar los procesos, más que evaluar el éxito alcanzado por personas o equipos de trabajo específicos.

En la evaluación de la conducta de los agentes, el paso siguiente es identificar la norma contra la cual las acciones desarrolladas para preparar el estudio ex-ante son evaluadas. Este es un punto extremadamente delicado, puesto que es escasa la legislación de orden general estableciendo pautas para este tipo de ejercicio retrospectivo. La norma, generalmente es establece en la forma de términos de referencia adosados a la contratación particular del agente. La participación de muchos agentes, no sólo dificulta la búsqueda de los estándares, también incrementa los costos de este tipo de estudios. ¿Son útiles, en este caso, las normas generales?. Por lo general, cada programa posee características propias que aconsejan preparar términos de referencia para cada caso. Sin embargo, es necesario considerar que las técnicas de manejo de información estadística que constituyen la práctica establecida avanza continuamente y no sería

de todo desacertado contar con alguna norma que en los términos de referencia para contratar estudios socioeconómicos ex-ante permita evaluar ese aspecto.

La revisión de la documentación existente debería brindar información que justifique porqué determinada dirección y no otra fue adoptada. Pero casi con certeza, no se encontrará documentación apropiada a este fin. Por esta razón es frecuente recurrir a entrevistas de participantes claves en la etapa de preparación previa del programa. Al redactar los resultados de las entrevistas debe tomarse en cuenta que los agentes tienden a omitir información sobre lo que podría haberse hecho pero no se hizo.

La norma contra la que se mide la bondad de la conducta de los agentes no necesariamente es de índole legal. Por ejemplo, podrían establecerse contenidos mínimos a cumplir, podría dejarse abierto al agente un rango acotado de opciones, podría emitirse juicio valorativo sobre las fuentes de información utilizadas o sobre los supuestos claves que guían el accionar del agente, etc. El aspecto que no puede perderse de vista es que el mayor interés, visto retrospectivamente, es identificar áreas de significativa divergencia a partir de algún indicador relevante. De esta manera, luego puede estudiarse las razones de la divergencia.

Estudios de revisión de confiabilidad

Con frecuencia, al término de los trabajos de análisis retrospectivo, de comportamiento de los agentes y de impacto, se incluyen recomendaciones orientadas a verificar la confiabilidad de la información suministrada, particularmente en lo que hace al análisis de impacto. El siguiente paso es, entonces, evaluar el grado de exactitud con que los estudios socioeconómicos anticiparon el impacto de decisiones administrativas arriesgadas durante la ejecución del programa, por ejemplo, un cambio de funcionarios, la transferencia del programa a otro ámbito institucional, modificación de los marcos regulatorios, o cambios en las restricciones presupuestarias y otras condiciones externas. Un estudio orientado a revisar la confiabilidad parte del mundo real para compararlo con las predicciones de impacto y pretende aislar las razones que dieron lugar a la existencia de inexactitudes, ponderando el impacto relativo de cada una de éstas.

En la evaluación retrospectiva, el concepto de confiabilidad en la información cubre tres aspectos: exactitud, calibración y consideración de la incertidumbre. La precisión., que se refiere al grado de correspondencia entre el impacto previsto ex-ante y el impacto efectivo, calibración, que apunta a clarificar los límites inferior y superior de la confiabilidad de un indicador de logros, y resolución, referida a la mayor o menor amplitud de las zonas grises que presentan los datos examinados y que vuelven incierto su mensaje. Desde luego, el orden de los capítulos en un estudio sobre la precisión de la información no se aparta de lo establecido. Pero razones de costo y efectividad sugiere que el mismo se focalice en aspectos particularmente problemáticos, o francamente descuidados y de alto impacto sobre las predicciones de resultados que en su momento justificaron el programa y hoy explicarían sus desvíos o su baja probabilidad de éxito. En un estudio de confiabilidad, es importante separar discrepancias debidas a falta de datos, o derivadas de supuestos o modelos de predicción incorrectos, de aquellas que surgen de una manipulación equivocada de la información disponible o un análisis deficiente de la misma por parte de los agentes involucrados en la ejecución. Estas últimas debieron haber sido identificadas como parte de los estudios sobre la conducta de los agentes. Un listado de identificación ayuda a identificar las fuentes de incertidumbre que deban ser exploradas en un estudio articular. En general, debe diferenciarse incertidumbres en los precios, en la selección de la tecnología (¿será aceptada por los usuarios? ¿y por los beneficiarios?. El servicio post-venta contratado, ¿será efectivo?), en la estimación del marco temporal durante el cual se sentirá el impacto del proyecto, en la selección de los indicadores apropiados para reflejar la fuerza del impacto, y en el valor atribuido socialmente al programa (a los supuestos e hipótesis iniciales que impulsaron la realización del programa, a la organización de la ejecución, al significado de los resultados alcanzados desde la perspectiva de diferentes actores).

5. Lecciones de la experiencia

Durante los últimos treinta años, la construcción de un pensamiento sistemático sobre evaluación retrospectiva de programas de mejora en la educación y otras áreas sociales registra un avance significativo.

Los indicadores de resultado han alcanzado una creciente aceptación que refleja un cambio actitudinal a favor de un ambiente de trabajo más competitivo dentro de las universidades públicas. Este fenómeno es común a universidades públicas europeas y de América Latina. Pero está claramente establecido que pocos indicadores serían de aplicación universal, por lo que su utilización más bien aparece como parte de un movimiento para asegurar una mayor calidad y control de los procesos a partir de ciertos resultados reconocidos y validados (Kells, 1993). En la Argentina, este mismo fenómeno ha comenzado a manifestarse en relación entre el gobierno nacional y las universidades públicas, por ejemplo, en relación a las carreras de riesgo (Morera y Brissón, 1998) mientras que las jurisdicciones provinciales registran una deuda con respecto a la evaluación retrospectiva de los colegios terciarios. Mejorar la comparabilidad significaría adoptar y hacer cumplir ciertos estándares en los métodos de cálculo, contar con normas de calidad sobre la preparación de la información y adoptar criterios comunes para incorporar la medición del riesgo y la incertidumbre.

El trabajo ha demostrado que es conveniente diferenciar tres modelos de evaluación retrospectiva, sea que se focalice en la conducción y ejecución del programa, los impactos, o la calidad de la información utilizada. Se señala, sin embargo, que los avances en el tratamiento de la incertidumbre son menores. Particularmente, no ha sido aún resuelto satisfactoriamente el tratamiento de la incertidumbre provocada por los cambios no anticipados en la dirección de la política, puesto que esto envuelve la tarea de ponderar el valor de una decisión en relación a otras posibles.

Referencias

Vani K. Borooah (1994), *Modelling Institutional Behavior: A microeconomic analysis of university management*, Public Choice 81:101-124.

José A. Delfino y Héctor R. Gertel (1996), *Financiamiento de las Universidades Nacionales: estímulo para mejorar su desempeño académico*, Actualidad Económica, Año VI, No 32 Mayo-Junio. pp.9-15.

Héctor R. Gertel (1984) Evaluación Ex-Post: Plan de desarrollo institucional de la Universidad Católica Madre y Maestra. Washington, Banco Interamericano de Desarrollo, Oficina del Contralor. Informe interno.

Héctor R. Gertel (1986) *Ex-Post Evaluation of Higher Education Development Programs in Latin America*. The XXX CIES Meeting at the Ontario Institute of Studies in Education, March 20/24.

Emanuel Jimenez y Yasuyuki Sawada (1999) *Do Community-Managed Schools Work?* The World Bank Economic Review, vol.13, No 3:415-442.

Jooseop Kim. Harold Alderman y Peter F. Orazem (1999) *Can private school subsidies increase schooling for the poor?* The World Bank Economic Review, vol.13, No 3:443-466.

Maurice Kogan (1993) *"The evaluation of Higher Education "*, en Maurice Kogan (Ed.) Evaluating Higher Education, Londres, Jessica Kingsley and OECD (pp. 11-24).

Henry M. Levin (1974) *"Education, Life Chances and the Courts "* Stanford University School of Education, mimeo.

Henry M. Levin (1983) *Cost-effectiveness: A Primer*. Beverly Hills, Sage Publications (New Series in Evaluation, vol.4).

Henry M. Levin (1997) *Raising school productivity: An X-efficiency Approach*, Economics of Education Review 16 (3):303-11.

María I. Morera y María E. Brissón (1998) *Carreras de Riesgo: El caso de Medicina. Metodología, procesos y productos en la formulación de los patrones y estándares previstos por la Ley de Educación Superior*. Asociación de Facultades de Ciencias Médicas de la República Argentina, Buenos Aires, Editorial La Prensa Médica Argentina.

OECD (1989) *Education and the economy in a Changing Society*, Paris, 1989

OECD (1999) *Redefining Tertiary Education*, Paris, OECD.

Alice Rivlin (1971) *Systematic Thinking for Social Action*, Nueva York, The Brooking Institution.

Gareth Williams (1994) *El financiamiento de la educación superior: Tendencias actuales*, México, Asociación Nacional de Universidades e Instituciones de Educación Superior.