

LOS RANKINGS COMO FORMA DE EVALUACIÓN

Dra. Hebe Irene Roig
Universidad de Buenos Aires

Agradezco mucho la invitación que me han hecho para participar en este encuentro tan importante para nuestras universidades. Vengo del campo de la pedagogía y mi interés es la evaluación como una posible herramienta de cambio y mejora de las instituciones educativas, de ahí la importancia de analizar el impacto de las diferentes formas de evaluación.

Como ustedes saben, la evaluación es una forma de intervención social, una forma de intervenir que trata de discriminar lo bueno de lo malo, lo mejor de lo peor, para orientar la toma de decisiones y las acciones consecuentes.

Existen diversas formas de evaluación, pero toda evaluación, según reconocen hoy la mayoría de los autores contemporáneos dedicados a este tema, no es sólo una práctica técnica, sino también política, ética e incluso, psicológica.

¿Cuál es el sentido de los rankings?, ¿es el marketing, es la evaluación? Su estrecha incidencia sobre la opinión pública y su canalización a través de los medios de comunicación, permitiría caracterizarlos como una forma de marketing que publicita las universidades mejor “rankeadas”, o mejor calificadas. Si acordamos con que la educación superior es un mercado, sería consecuente pensar los rankings en términos de marketing.

Si en cambio, creemos que la educación superior es un derecho social y no un mercado, debemos analizar los rankings como forma de evaluación educativa.

Desde este segundo punto de vista tenemos la expectativa de que los rankings sean un estímulo para la mejora de la calidad de las universidades hacia su misión institucional.

Voy a recuperar algunos conceptos clásicos de la evaluación, conceptos que se discutieron mucho en el siglo pasado, particularmente desde la década de los 60`s en el mundo anglosajón. Estos conceptos hoy se asumen como fundamentos del campo teórico de la evaluación.

Las dos principales funciones de la evaluación son la rendición de cuentas y la de mejorar. Rendir cuentas ante la sociedad y buscar caminos para la mejora hacia el interior de las instituciones.

Algunos tipos de evaluación cumplen mejor que otras estas funciones. Una distinción clásica es la que hizo Michael Scriven en los 60's, distinguiendo evaluación de resultados de evaluación de procesos en la calificación de los estudiantes.

La evaluación de resultados califica solamente el resultado final, pero no da elementos para comprender por qué se producen esos resultados. Los rankings son una evaluación de resultados.

En las evaluaciones basadas en indicadores, datos descriptivos de un momento institucional, también se dejan de comprender los procesos.

Los contextos y sus condiciones, los juegos de poder y los saberes, valores, creencias, actitudes, roles y relaciones de las culturas institucionales no pueden expresarse en la evaluación de resultados. Para los rankings, las universidades son una “caja negra”.

En contraposición, las formas de evaluación que atienden a los procesos, más allá de medir, recurren a otras metodologías de investigación.

La evaluación de procesos, sí permite el aprendizaje y la mejora, porque se orienta a comprender más que a medir. Si el objetivo de los rankings es que también las autoridades puedan tomar decisiones, hay que tener presente que esta forma de evaluación no da pistas para comprender el por qué de los problemas. Sí da elementos para detectar puntos débiles, pero sólo los puntos débiles jerarquizados por quienes determinan los indicadores.

La evaluación no es sólo medir, es emitir un juicio basado en una comparación. Las evaluaciones también han sido descritas tomando el criterio de qué se compara con qué. Aquí cabe distinguir entre evaluación preordenada y evaluación constructivista. Evaluación preordenada es un concepto de Robert Stake (1967). Esta forma evaluativa define a priori qué va a ser evaluado y cómo va a ser medido. Implica un modelo de universidad previamente definido.

Un impacto característico de esta forma de evaluación es que lo que se evalúa se convierte en importante. Pedro Krostch (2007) investigador de la educación superior en Argentina, analizó el impacto de las evaluaciones institucionales de la CONEAU luego de 10 años de actividad. Observa que, en consecuencia con los indicadores priorizados, el mayor impacto radicó en la revisión de normativa interna y en ciertos ajustes a los sistemas administrativos universitarios.

Hoy las evaluaciones preordenadas que más utilizamos son los sistemas de indicadores y la evaluación a través de estándares.

La evaluación preordenada parte del supuesto de que conocemos de antemano cómo deben ser las cosas. Los criterios de comparación se construyen antes de observar los fenómenos. En este enfoque hay variantes, los objetivos iniciales pueden ser considerados como los puntos de referencia para la evaluación, pero también el conocimiento experto. En este último caso, un experto, o un equipo de expertos sobre determinada temática universitaria, construyen una visión modélica que servirá de parámetro a la comparación.

Por ejemplo, los sistemas de indicadores para evaluar se basan en un cuerpo teórico que identifica cuáles son las variables e indicadores relevantes que inciden en el fenómeno objeto de estudio. El estudio de la presencia de esos indicadores da elementos para posteriormente interpretar la información y construir un juicio sobre lo que se observa.

Los indicadores ofrecen información sintética, una especie de radiografía, que puede ser insumo para debatir y asumir decisiones. Permiten establecer comparaciones con otros fenómenos u objetos de estudio semejantes, y la medición continua de los mismos indicadores permite estudiar tendencias en el fenómeno. Por último, facilitan el debate público porque son comprensibles para los ciudadanos.

Pero también, son un instrumento con limitaciones.

Cabe aquí una reflexión de Broadfoot (2002), un investigador del campo educativo: “el primer paso es medir aquello que se puede medir con facilidad. Hasta aquí todo va bien. El segundo paso es no tener en cuenta lo que no se puede medir fácilmente o adjudicarle un valor cuantitativo arbitrario. Esto es artificial y engañoso. El tercer paso es asumir que lo que no se puede medir fácilmente no es de verdad importante. Esto es ceguera. El cuarto paso es decir que lo que no se puede medir con facilidad en realidad, no existe. Esto es suicida”.

En tanto que sólo transmiten una idea sintética, pueden interpretarse erróneamente relaciones causales entre ellos cuando son solamente mediciones descriptivas. Asimismo, pueden tener también diferentes interpretaciones y por lo tanto, derivarse de ellos líneas de acción diferentes.

En la práctica concreta de la evaluación, los indicadores no son solamente una construcción teórica. Si se mide un indicador es porque se lo ha considerado relevante, tal como dice Tiana Ferrer (1997), “la construcción de un indicador (...) es el resultado de una compleja interacción entre un proceso cognitivo y otro político”.

Otra forma de construir la comparación que da base al juicio evaluativo es la evaluación constructivista, concepto acuñado por Guba y Lincoln (1989). Aquí, los términos de comparación no se definen antes, sino que son construidos en el propio proceso de evaluación.

En este caso, la evaluación deberá descubrir y jerarquizar los problemas que enfrentan los actores de una institución, para luego buscar las mejores formas de comprenderlos y proponer alternativas de acción.

Mientras esta segunda forma se adapta a contextos y características particulares, la evaluación preordenada se focaliza en aspectos generales, tal como es el caso de los rankings.

Otro aspecto que considero relevante incorporar a estas reflexiones, es cómo definimos la validez de la evaluación.

Ernest House (1994) sostiene que la validez de la evaluación depende de tres criterios: ser veraz, creíble y ajustada al derecho. La veracidad depende de la rigurosidad de los instrumentos, de la validez metodológica y la confiabilidad en los métodos de investigación para relevar información. La credibilidad, depende de cómo establece la comunicación con los destinatarios, cómo plantea los argumentos y los persuade de su valor. Por último, que sea ajustada al derecho alude al respeto a los derechos de los actores involucrados y a la ética.

Esta manera de considerar la validez de la evaluación hace énfasis en que no hay que confundir validez metodológica con validez de la evaluación. La evaluación válida es aquella que cumple con los criterios de credibilidad, derecho y ética. Por tanto, la evaluación no puede desentenderse de sus consecuencias.

Es parte de la evaluación asumir la responsabilidad sobre las consecuencias. Los productores de ranking no pueden argumentar la ignorancia de los medios o de la opinión pública. La ética en la evaluación reclama asumir la responsabilidad sobre los usos de la información. Si los rankings son malinterpretados por el sentido común de los medios y la opinión pública, no es un problema del público, es un problema de comunicación de los resultados.

Las decisiones no se toman sobre la información de modo directo, sino sobre cómo se interpreta dicha información en un determinado contexto de uso. Sabemos la diferencia entre una interpretación dada desde el sentido común y la que se establece cuando se cuenta con una mirada formada en un campo disciplinar. Autoridades desprevenidas sobre las características metodológicas de los rankings pueden también malinterpretar su sentido y caer en la búsqueda de modificar el valor de un indicador, en vez de mejorar la calidad de su institución.

Los rankings y clasificaciones de instituciones basados en medición de indicadores no hacen transparentes a las universidades, no son una fotografía de las universidades, son una radiografía. Son una herramienta técnica de difícil interpretación correcta por legos. Si se dice que se hace investigación comparada mediante un sistema de indicadores, corresponde hacer divulgación científica. Si se dice que se hace evaluación, se debe alertar sobre los límites del propio modelo evaluativo.

La validez de consecuencia, nos lleva nuevamente a preguntarnos por cuál es el impacto de las diferentes formas de evaluación que existen y cuál es la característica de los rankings. Frente a una forma evaluativa que alimenta la competencia como principio de la posible mejora de las instituciones, otras formas evaluativas fortalecen el trabajo en equipo y la autosuperación. La metáfora que podría ejemplificar esta diferencia es una carrera de cien metros frente a otro deporte, que también implica competencia, pero pone en juego otras reglas: el montañismo. En esta última actividad deportiva la vida depende del trabajo en equipo y la solidaridad de todos sus miembros, y a la vez, la superación de las propias marcas es el principal objetivo.

La mejora universitaria no es cambiar el valor de un indicador, es lograr modificar condiciones, prácticas y culturas institucionales que aumenten la posibilidad de dar mejores respuestas frente a su misión institucional. Los sistemas de ranking no son instrumentos de mejora para las universidades, pueden servir al debate de políticas macro sobre algunos aspectos puntuales, pero no son suficientes para el desarrollo de nuestras universidades. Se puede, por ejemplo, decidir el aumento de presupuesto para ciertas áreas, pero como ustedes saben, contar con más recursos no significa necesariamente el aumento de la calidad porque se necesita también resolver problemáticas internas que dependen de las dinámicas institucionales.

La evaluación como una posible herramienta de cambio y mejora de las instituciones educativas sigue siendo un desafío. Creemos, más allá de la competencia, en la necesaria diversidad, en la convergencia de esfuerzos para la resolución de problemas comunes tanto del desarrollo institucional como de los contextos sociales, económicos y culturales a los que pertenecemos.

México D.F., mayo de 2012.

REFERENCIAS BIBLIOGRAFICAS

- Guba, E. G. y Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park, Ca.: Sage Publications.
- House, E. (1994) *Evaluación, ética y poder*. Madrid: Morata.
- House, E. (2001) *Valores en la evaluación e investigación social*. Madrid: Ediciones Morata.

- Krotsch, P. y otros (2007). *Evaluando la evaluación. Políticas universitarias, instituciones y actores en Argentina y América Latina*. Buenos Aires: Prometeo.
- Kushner, Saville. (2002). *Personalizar la evaluación*. Madrid: Morata-Fundación Paideia Galiza.
- P. Broadfoot, en T. Atkinson y G. Claxton, 2002. *El profesor intuitivo*. Barcelona: Octaedro.
- Stake, R. (2006). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Graó.
- Stufflebeam, D. y Shinkfield, A. (1987) *Evaluación sistemática*. Barcelona: Paidós.
- Tiana Ferrer, Alejandro “La evaluación de los sistemas educativos”. Revista Iberoamericana de Educación. OEI. Número 10, enero-abril, 1996. Consulta en Internet: septiembre de 2011. <http://www.rieoei.org/oeivirt/rie10a02.pdf>