

El modelo de evaluación de carreras: más de lo mismo

Arturo Villavicencio

Universidad Andina Simón Bolívar

Junio 2014

Desde los años 1990 en México y otros países latinoamericanos se implantó un aparato de evaluación sustentado en la promesa de la calidad. Decenas de millones de niños y jóvenes mexicanos fueron medidos con pruebas estandarizadas, cientos de miles de maestros y académicos sujetos a evaluación para merecer un pago adicional y decenas de miles de escuelas, programas y universidades colocadas bajo escrutinio permanente. La evaluación 'moderna' tuvo éxito en crear un enorme aparato de vigilancia dedicado a medir a la nación una y otra vez, pero fue incapaz de generar una dinámica de mejoramiento de la educación y más bien contribuyó a degradarla (Aboites, 2012).

Resumen

La Universidad ecuatoriana se apresta a ser evaluada nuevamente. Esta vez son las carreras de medicina el objeto de un examen y escrutinio. Siguiendo los lineamientos de un 'modelo genérico de evaluación de carreras', la entidad evaluadora ha elaborado un modelo específico que repite las mismas deficiencias y errores conceptuales y metodológicos del modelo aplicado en la última evaluación de universidades. La estructura de evaluación a ser aplicada contraviene principios básicos en los que se sustentan las herramientas de tipo multicriterial. La redundancia de los criterios e indicadores, la irrelevancia de algunos para los objetivos de la evaluación y la preocupante ambigüedad en su definición son manifestaciones de serias omisiones de las condiciones de exclusividad, pertinencia y operatividad que deben cumplir el conjunto de criterios e indicadores. Además, el modelo no explicita los arbitrajes necesarios para introducir ajustes de compensación entre criterios y/o indicadores, así como para resolver situaciones de proximidad a límites o marcas establecidas. Una revisión del modelo resulta indispensable para asegurar niveles aceptables de integridad y legitimidad en el proceso de evaluación. La revisión, necesariamente debe partir de un diálogo y debate con todos los actores involucrados.

1. Introducción

Un aporte trascendente del Mandato Constituyente No.14 fue el de haber reposicionado la evaluación universitaria como un mecanismo de garantía pública del funcionamiento de las Instituciones de Educación Superior (IES) dentro de un marco legal de autorización, reconocimiento, validez y certificación de sus actividades. Luego de una corta experiencia de una evaluación formal y complaciente, limitada al recuento de indicadores ambiguos y dudosos, que en definitiva no indicaban nada, la evaluación bajo el Mandato Constituyente generó confianza y expectativas sobre la recuperación del sistema universitario y sirvió para legitimar social y académicamente la presencia de una institución con la misión de impulsar el mejoramiento permanente de la educación superior. Quedó demostrado que era posible empezar la construcción institucional con la capacidad resolver las inherentes tensiones entre universidad,

mercado y gobierno como mecanismo de asegurar un proceso continuo de avance en la calidad de la educación universitaria.

La evaluación realizada el año anterior por la agencia evaluadora pone en riesgo la realización de esas expectativas. En lugar de emprender e impulsar un proceso de largo alcance de innovación y mejora de los sistemas educativos, la universidad ecuatoriana ha entrado en la cultura de la evaluación, entendida como un mecanismo inquisidor, de vigilancia y control, que la somete a los dictámenes de estructuras burocráticas de poder. Se está allanando el camino para que las universidades, de instituciones relativamente autónomas pasen a ser simples dependencias de los órganos de regulación del sistema de educación superior (si es que en realidad ya no lo son), controladas por el Gobierno y por tanto sujetas a su supervisión y vigilancia. En este contexto, la evaluación viene a ser un símbolo de la burocratización y subordinación que desdibuja la universidad como factor decisivo en la vida nacional. De hecho, se crean las bases de un nuevo esquema corporativo entre el Estado y las instituciones de educación superior, estas últimas como productoras de servicios y conocimiento y el primero como regulador de la competitividad y el mercado.

Toda esta política de control pretende ser legitimada bajo una supuesta neutralidad y objetividad de la 'metodología' de evaluación. Una serie de conceptos y herramientas, cuya aplicación tuvo validez y pertinencia en el manejo de la complejidad del cumplimiento del Mandato Constituyente 14, son presentadas ahora como instrumentos de empleo universal para la solución de problemas que van desde la optimización de recursos y energías renovables (CEAACES, 2013, pág. 11), hasta el estudio de las galaxias en astronomía (CEAACES, 2013, pág. 13), pasando, por supuesto, por la evaluación de la calidad de sistemas de educación superior. Así, la presunta científicidad del ejercicio de evaluación se expresa mediante el uso de fórmulas inescrutables, índices nanométricos, rangos de confiabilidad arbitrarios.

La universidad, entonces, soporta una engorrosa tarea de recolección de datos y verificación de indicadores bajo la práctica de un empirismo que pretende convencerla que los datos constituyen una "descripción científica y objetiva de la realidad" y que las inferencias que se desprenden de ellos son legítimas para sustentar las políticas universitarias, cuya validez no admiten cuestionamiento. Al respecto, conviene anotar que:

"el debate epistemológico acerca de la calidad científica de la evaluación adquiere actualidad porque el pensamiento dominante requiere que el empirismo recupere el prestigio perdido en la historia de la ciencia y para ello echa mano de aportes metodológicos reconocidos en los círculos académicos. El enfoque empirista de la evaluación de la educación se sustenta en las metodologías neopositivistas. Los centros de pensamiento pretenden conferirle el estatus de alta ciencia a la evaluación tecnicista por medio de sutiles mecanismos de legitimación" (Leher, 2012, pág. 26).

Estas pretensiones de validez del discurso científicista de una evaluación tecnicista han sido objeto de críticas radicales que lo dejan sin sustento alguno. Al demostrar Aboites (2012) que este tipo de evaluación carece de neutralidad y, por lo tanto, las razones de su difundida práctica no responden a preocupaciones pedagógicas ni de mejoramiento de la enseñanza, el debate se

plantea entonces en su debido lugar: la política. Esta es la razón por la cual el tema de la evaluación universitaria debe formar parte de la agenda sobre el debate de la reforma del sistema de educación superior; debate hasta hoy negado a la comunidad universitaria.

Los procesos de evaluación deben avanzar y consolidarse en el sistema de educación superior. Pero hay que entender al tipo de evaluación al que nos estamos refiriendo. No se trata de establecer una cultura de fiscalización y obediencia, reducida a la concepción y aplicación de un conjunto de herramientas y procedimientos ‘técnicos’ que pretenden suplir la falta de ideas y conceptos bajo una pantalla de objetividad aritmética. Se requiere, ante todo, una evaluación para la acreditación universitaria; es decir, una certificación para la sociedad de que las instituciones de educación superior operan con licitud y legalidad, al amparo de una garantía pública sobre el alcance de sus actividades: autorización, reconocimiento, habilitación, titulación, credencial y validez (Dias Sobrinho J. , 2008 b). Las condiciones están dadas para ello.

2. Criterios e indicadores en duda

Las ciencias sociales distinguen dos enfoques o métodos para la aprehensión de la realidad social (Abbott, 2001; Grin, Rotmans, & Schot, 2010). El uno asume que el mundo social consiste de entidades fijas (las unidades de análisis), que poseen atributos (las variables). Estas variables interactúan para dar resultados que son medibles como atributos de las entidades; en otras palabras, los resultados son el producto de variables independientes que actúan sobre variables dependientes. La literatura se refiere a este enfoque como el *modelo lineal general o la teoría de la varianza*. El segundo enfoque, la *teoría de procesos o narrativas*, se centra en eventos antes que variables causales. Los eventos son producidos por actores, quienes toman decisiones, emprenden acciones y reaccionan entre ellos. Ellos identifican patrones y mecanismos sobre la base de los resultados de secuencias temporales de eventos.

La evaluación de las instituciones de educación superior con fines de acreditación corresponde a un caso especial del modelo lineal general. Se trata de un análisis de la situación de la universidad en un momento dado. Bajo esta perspectiva, el problema de análisis puede ser enfocado como un conjunto de entidades, las instituciones de educación superior, con una identidad unitaria, dotadas un conjunto de atributos que reflejan su desempeño sincrónico frente a ciertas propiedades o variables. El universo de evaluación es ‘variabilizado’ (Grin, Rotmans, & Schot, 2010); es decir, consistente de variables o dimensiones interrelacionadas. El análisis se sitúa en dos niveles (Abbott, 2001). Al nivel conceptual, la relación de causalidad entre criterios es de carácter sintáctico (varios conceptos operan en un orden que configura la estructura de evaluación), pero al mismo tiempo, la relación entre conceptos e indicadores es también semántica (los indicadores apuntan a conceptos difusos que tratan de representarlos en una forma práctica y operacional).

Esta dualidad da lugar a diferentes categorías de ambigüedad: ambigüedad semántica (un indicador significa más de un concepto); ambigüedad sintáctica (un indicador significa un concepto con diferente explicación de causalidad); ambigüedad contextual (significado contextual de los indicadores); y ambigüedad interactiva (el significado de un indicador depende de la intensidad de

interacción con otros indicadores) derivada del grado de compensación entre indicadores. Es en este contexto de incertidumbre que el ejercicio de evaluación puede ser planteado como un problema de carácter multicriterial, cuyos conceptos y herramientas permiten manejar la complejidad inherente a este tipo de problemas.

Este enfoque exige que las variables o criterios cumplan, al menos, tres condiciones fundamentales (Harguindéguy, 2013) :

- a) ante todo, las variables (criterios, indicadores) deben ser ‘conjuntamente exhaustivos, es decir, deben cubrir el conjunto de los atributos del fenómeno que se analiza;
- b) las variables deben ser ‘mutuamente exclusivas’, es decir, deben enfatizar un aspecto del fenómeno sin solaparse con los demás; y
- c) las variables deben tener significado para todos los actores involucrados en el ejercicio de evaluación. Esta condición supone: i) claridad de los criterios e indicadores respecto a sus implicaciones sobre los objetivos que se buscan al evaluar las entidades; ii) relevancia de las variables sobre los objetivos; y iii) transparencia y claridad en el contenido y alcance de las variables con el fin de facilitar el diálogo entre las partes que intervienen en el proceso.

Al igual que el modelo de evaluación de las universidades aplicado por el CEAACES, el modelo de evaluación de la carrera de medicina adolece de serias deficiencias respecto al cumplimiento de las condiciones básicas requeridas para la integridad de un modelo de evaluación. Varios ejemplos a continuación corroboran esta afirmación.

Redundancia

Existe una redundancia evidente en el alcance y significado de varios indicadores del modelo de evaluación de carreras (CEAACES, Junio, 2014). Esta reiteración de los indicadores conduce a centrar la atención en determinados aspectos o propiedades de los objetos de evaluación que, en definitiva, se traduce en una sobre ponderación de determinado atributo o cualidad. Así, una universidad con un desempeño notable respecto a un atributo sobre el cual existe redundancia en los indicadores se verá netamente favorecida en comparación con una universidad con un desempeño inferior respecto al atributo en cuestión. El Cuadro 1 presenta algunos ejemplos de redundancia de varios indicadores del modelo.

Este énfasis reiterativo en ciertos atributos o propiedades que se desea evaluar vulnera el principio de exclusividad de criterios e indicadores e introduce un sesgo evidente en la evaluación, centrándose en aspectos que muchas veces son seleccionados por la facilidad en el manejo de la información existente o potencialmente disponible (evidencias). Esta tendencia conduce a privilegiar la información fácilmente disponible y centrada en insumos, bajo la hipótesis que se trata de información independiente (no redundante), cuando en realidad existe un número limitado de elementos interactivos entre la mayoría de ellos¹.

¹ Al respecto, cabe señalar que en la evaluación de las IES (CEAACES, 2013), un número importante de indicadores no expresaban sino la disponibilidad de recursos económicos de las universidades.

Cuadro 1. Ejemplos de redundancia en los indicadores

| | | |
|---|---|---|
| Perfil profesional consultado (C.1.1) | ↔ | Estado actual y prospectiva (P.1.1) |
| La carrera debe demostrar que el Perfil Profesional Consultado ha sido establecido en base a estudios y análisis de las necesidades del entorno (estado del arte de las áreas relacionadas con la carrera, de la planificación nacional o regional, necesidades del sector productivo, criterios de los empleadores y de los graduados, entre otros). | | Análisis de pertinencia de la carrera. Se debe mostrar, de qué manera el diseño de la carrera responde: a) Las necesidades del entorno provincial y local; b) Las necesidades del entorno nacional; c) Las necesidades del desarrollo científico-tecnológico; d) Los requerimientos de la planificación nacional y regional |
| Bibliografía básica (I.1.1) | ↔ | Calidad bibliográfica (I.1.2) |
| El indicador mide la disponibilidad promedio de material bibliográfico para cubrir las necesidades básicas (bibliografía obligatoria) de las asignaturas de la carrera. Se tomará una muestra de diez asignaturas de la malla. | | El indicador evalúa la calidad bibliográfica de la carrera, a través de los siguientes criterios generales: 1. Pertinencia, 2. Acervo (clásico y contemporáneo, físico y virtual), 3. Conservación, renovación, 4. Presupuesto ejecutado en material bibliográfico. |
| Seguimiento a graduados (P.1.2) | ↔ | Perfil de egreso |
| Evalúa la existencia, funcionamiento y conocimiento de la comunidad académica [sic]; lo que permite el monitoreo periódico de la evolución profesional de los graduados de la carrera, con el objetivo de identificar las posibles mejoras que se puedan introducir en el Perfil de Egreso y Currículo. | | Evalúa la coherencia entre el Perfil de Egreso y el Perfil Profesional en relación a la pertinencia de la carrera. |

Congruencia

Varios indicadores se refieren a disposiciones establecidas explícitamente en el Reglamento de Régimen Académico y en el Reglamento de Carrera y Escalafón de la Docencia. En este caso, el indicador debe ser congruente con lo estipulado en los reglamentos, especialmente en lo referente a sus rangos de variación dentro de los cuales la valoración de desempeño es 'óptima'. Esta observación es muy pertinente, considerando que en la evaluación de las universidades realizada el año anterior (CEAACES, 2013), el modelo aplicado definió de manera arbitraria los parámetros de valoración; contradiciendo claras disposiciones reglamentarias. El cuadro 2, a continuación, destaca esta situación para los indicadores sobre la dedicación efectiva de docencia. El modelo de carreras propuesto incluye tres indicadores sobre la carga horaria de los docentes y, aunque no es explícito sobre los criterios de valoración que se utilizaran para estos indicadores, resulta oportuno destacar esta incongruencia para evitar repeticiones.

Una observación similar es válida para el indicador "Docentes tiempo completo" (A.2.1). Mientras el reglamento de la LOES (Transitoria 14) claramente estipula que al menos el 60% de la planta docente de una universidad debe consistir de profesores a tiempo completo, el modelo aplicado por el CEAACES adoptó para este indicador una función de valoración de la utilidad que asignaba la máxima puntuación para un intervalo a partir del 80 por ciento.

**Cuadro 2. Carga horaria de los docentes:
evaluación versus reglamentos**

| Indicador | Carga horaria de los docentes (horas de clase por semana) | |
|---------------------------------|--|-----------------------------|
| | Modelo evaluación IES | Reglamento Escalafón |
| Tiempo completo (A.2.3) | 6.25 – 18.75 | 3 - 16 |
| Docentes medio tiempo (A.2.4) | 4 - 12 | 10 |
| Docentes tiempo parcial (A.2.5) | 4 - 12 | 2 - 9 |

De igual manera, el indicador “Actualización científica” (A.1.3) “evalúa la participación de los docentes en eventos de actualización científica realizados en el último año anterior a la evaluación, tales como seminarios, talleres, cursos y otros, en el área en la que el docente ejerce la cátedra”. Al respecto, el Reglamento de escalafón (art. 53 – 57), entre los requisitos para ascenso entre las diferentes categorías de docencia, establece claramente los requisitos de ‘actualización científico – profesional’ para los docentes’ según se resume en el Cuadro 3 a continuación.

Cuadro 3. Requisitos de actualización profesional y científica de los docentes

| De: | A: | Horas |
|------------|------------|--------------|
| Auxiliar 1 | Auxiliar 2 | 90 |
| Auxiliar 2 | Agregado 1 | 180 |
| Agregado 1 | Agregado 2 | 300 |
| Agregado 2 | Agregado 3 | 400 |
| Agregado 3 | Principal | 480 |

Otro ejemplo está dado por el indicador “Seguimiento a graduados” que evalúa la existencia de un “proceso que permita monitorear periódicamente la evolución profesional de los graduados de la carrera...”. Sin embargo, el Reglamento a la LOES establece que “la SENESCYT diseñará los procedimientos necesarios para que las instituciones de educación superior instrumenten un sistema de seguimiento a los graduados, el cual será parte del SNIESE” (art. 26). Entonces, resulta un tanto incomprensible verificar y calificar el cumplimiento de una disposición que todavía no ha sido normada por la institución responsable, como lo dispone el Reglamento.

En general, para los indicadores que se refieren a normas explícitas de reglamentos y normas que rigen el funcionamiento del sistema de educación superior, carece de sentido valorar el desempeño de las IES mediante una función definida sobre la base del concepto de utilidad. ¿Qué significado puede tener afirmar que una IES cumple satisfactoriamente, parcialmente o de manera deficiente tal cual norma de un reglamento? El problema consiste en una confusión entre evaluación para la acreditación y la puesta en marcha de programas para el mejoramiento la enseñanza. La primera se refiere a un proceso de verificación y garantía a la sociedad de que las instituciones de educación superior operan dentro de los márgenes legales definidos por el conjunto de leyes, normas y reglamentos que regulan su funcionamiento. Este constituye, además, el mecanismo apropiado para evitar la repetición de experiencias nefastas del pasado, donde la anomia institucional puso en juego la legitimidad misma de la universidad. Es probable que este enfoque de la evaluación requiera la introducción de estándares y criterios adicionales. Si tal es el caso, estos deben sustentarse en un proyecto claro de mejora de la enseñanza, consistentes con normas y reglamentos y, sobre todo, consensuado con los actores del sistema y no impuesto unilateralmente por instancias burocráticas.

Pertinencia

Una abundante literatura ha sido dedicada al análisis del impacto real del entorno sobre la efectividad del aprendizaje (Dill & Soo, 2005; Astin, 1996; Terenzini & Pascarella, 1994). Estudios empíricos (Terenzini & Pascarella, 1994) muestran que la mayoría de los indicadores comúnmente utilizados en las evaluaciones tienen un efecto mínimo o irrelevante en el aprendizaje de los estudiantes, convirtiendo estos indicadores en mitos de la educación superior (Terenzini & Pascarella, 1994). El efecto de varios indicadores sobre la ‘calidad’ de la enseñanza es parte de aquella “sabiduría convencional” sobre la educación superior; un conjunto de ideas que descansan en hipótesis más o menos ingenuas, no comprobadas y simplificadoras de la realidad universitaria que hacen ver su problemática fuera de su contexto específico. Son estas hipótesis las que llevan a construir el problema en una dirección específica e, implícitamente, conllevan a una solución predeterminada que legitima las políticas y acción de las autoridades. Un análisis detallado sobre este tema está fuera del objetivo del presente trabajo; sin embargo, resulta de interés detenerse en examinar el alcance y significado de tres criterios o indicadores como un ejemplo de ideas convencionales que resultan contra intuitivas en la evaluación de las universidades. Estos criterios son: la investigación como uno de los factores determinantes de la calidad de la enseñanza; el ‘nivel óptimo’ de la relación estudiantes – docentes y la eficiencia terminal medida como el porcentaje de graduados en el ‘tiempo reglamentario’.

- ***Investigación***

El modelo de evaluación de universidades (CEAACES, 2013) asignó el veinte por ciento de la evaluación de una universidad a su desempeño en investigación, medida básicamente por el número de publicaciones: artículos publicados en las revistas que figuran en las listas de Scopus o ISI, el número de artículos publicados en la lista de revistas de Latindex (bajo la curiosa denominación de ‘investigación regional’) y el número de libros revisados por pares. En El modelo genérico de evaluación de carreras y el modelo de evaluación de carreras de medicina incluyen

esta concepción de la investigación como criterio de la evaluación². Dada la importancia asignada a este criterio en la última evaluación, es muy probable que su peso sea así mismo decisivo en la nueva evaluación.

La importancia del criterio investigación (medido por el número de publicaciones en revistas indexadas) se sustenta en el supuesto según el cual existe una relación directa entre la intensidad de la investigación y la calidad de la enseñanza; es decir, mientras mayor sea el número de publicaciones de los docentes de una universidad, mejor será su calidad de enseñanza. Según este razonamiento, una universidad dedicada a la investigación, calificada como tal según el número de artículos publicados por sus docentes, contribuye al desarrollo de un país en una escala mayor que una universidad cuyos docentes publican menos. Este supuesto, si no totalmente errado, por lo menos resulta muy cuestionable.

Este cuestionamiento conduce al planteamiento de la pregunta: ¿las universidades con una fuerte orientación hacia la investigación ofrecen el mejor entorno de aprendizaje y proveen la mejor educación? La respuesta requiere ante todo aclarar el significado de la expresión ‘universidad orientada a la investigación’. Astin (1996) la define en términos de comportamiento y actitud: “el volumen de publicaciones de los docentes, el tiempo que ellos invierten en la investigación y su compromiso personal con la investigación”. Entonces, un docente con una marcada orientación a la investigación es aquel que publica muchos artículos y libros, dedica la mayor parte de su tiempo de trabajo a la investigación y asigna una prioridad personal a esta actividad. Como antítesis a esta categoría de universidad, el autor mencionado distingue una ‘universidad orientada a los estudiantes’ (o universidad de docencia) como aquella cuyos objetivos son “enseñar a los estudiantes como cambiar la sociedad, desarrollar liderazgo entre los estudiantes, inducir a la reflexión sobre problemas sociales y ambientales, ayudar a los estudiantes a entender sus propios valores, desarrollar un sentido de comunidad entre docentes y estudiantes y facilitar el involucramiento de los estudiantes en actividades y servicios sociales y comunitarios” (Astin, 1996). La pregunta anterior puede ser entonces reformulada en los siguientes términos: ¿Qué tipo de universidad, la orientada a la investigación o la orientada a la docencia, es preferible en términos de un mejor aprendizaje y educación?

Los resultados de un volumen significativo de investigación empírica sugieren que la correlación entre la productividad de la investigación y la educación de pregrado es negativa o, en el mejor de los casos insignificante y que la enseñanza y la investigación parecen ser más actividades independientes que actividades articuladas entre sí (Astin, 1996; Dill & Soo, 2005; Donoghue, 2008)³. Los autores mencionados coinciden en afirmar que la orientación de la docencia hacia la investigación incide negativamente en una serie de factores que tienen que ver con la enseñanza: horas dedicadas a la enseñanza y consulta, compromiso con el desarrollo de los

² Una discusión sobre el tema de las ‘publicaciones científicas’ se puede encontrar en: Villavicencio, A. (2014), *La fiebre de publicaciones*; <http://lalineadefuego.info/category/educacion/>

³ Donoghue (2008, pág. 51) menciona que de una revisión de treinta estudios sobre la relación entre investigación y enseñanza en universidades estadounidenses, un tercio encontraron una correlación positiva muy pequeña, uno encontró una correlación negativa y dos tercios no encontraron ninguna correlación.

estudiantes, uso de técnicas activas de aprendizaje en el aula y el porcentaje de docentes comprometidos con un sentido de comunidad universitaria. La explicación es simple: los docentes dedicados a la producción de 'papers' pueden dedicar únicamente un tiempo muy limitado de su tiempo académico a las actividades mencionadas. Los autores mencionados concluyen que la orientación hacia la investigación tiene efectos negativos en la satisfacción de los estudiantes con la docencia⁴. "Existe in conflicto obvio entre investigación [en el sentido anteriormente definido] y docencia" (Astin, 1996, p. 592).

- Estudiantes y docencia

Debido a la simplicidad y existencia de la información, la relación estudiantes/docentes es uno de los indicadores privilegiados para medir la calidad de la enseñanza. La lógica de este indicador es bastante simple: mientras menor es el número de estudiantes por docente, mayor es la atención que los profesores pueden ofrecer a sus estudiantes e inversamente, mientras más elevada es esta relación, menor es el tiempo disponible de los docentes para los estudiantes. Schenker e Inauen (2011) señalan la ausencia de un estudio, empíricamente validado, sobre un estándar adecuado para este indicador. Anotan que la mayoría de los estudios realizados se refieren a la educación primaria y secundaria y los resultados no pueden ser aplicados a la educación superior.

Resulta muy difícil argumentar que el tamaño de una clase captura las dimensiones críticas del currículo y la enseñanza que las universidades pueden usar para mejorar la calidad de la enseñanza y el aprendizaje (Dill & Soo, 2005). No existe un estándar óptimo de la relación estudiantes/docente para todas las disciplinas y tipos de universidades⁵. Cada disciplina tiene su currículo estándar específico y cada universidad educa a sus estudiantes bajo diferentes modelos de enseñanza y, en ambos casos el nivel de intensidad en tiempos de dedicación es diferente. Así, por ejemplo, en algunas disciplinas puede dominar la modalidad de conferencias magistrales, en otras son más importantes los seminarios y clases prácticas, otras requieren más horas de laboratorio. En cada caso, el tamaño 'óptimo' de cada clase y el tiempo de dedicación varía, así como varía la categoría de la docencia pertinente.

⁴ Estas observaciones son suficientes para llamar a una seria reflexión en los estándares de evaluación definidos por el CEAACES en la evaluación de las IES (CEAACES, 2013). De acuerdo a estos estándares (benchmarks) la producción ideal de un docente a tiempo completo es como sigue: 1 libro cada tres años, 1.5 artículos por año publicados en las revistas que figuran en las listas Scimago (Scopus) o ISI Web of Knowledge y 2 artículos por año en una de las revistas de Latindex. En otras palabras, se espera que a lo largo de una carrera docente de treinta años, un profesor universitario haya publicado por lo menos 10 libros y más de cien artículos en revistas científicas. Esto puede traer dificultades en el futuro ya que con ese ritmo de publicaciones, las revistas científicas internacionales van a estar saturadas de artículos de académicos ecuatorianos. Una situación similar se puede producir con la industria editorial del país, cuya capacidad puede resultar insuficiente para cubrir la demanda de publicación de libros de los docentes universitarios del país.

⁵ La Conferencia de Rectores de las Universidades Suizas (CRUS) recomienda para el nivel de maestría un relación estudiantes-docentes de 40:1 en humanidades y ciencias sociales, 35:1 en carreras técnicas y una relación 60:1 para el nivel de pregrado en todas las disciplinas (Schenker-Wicki & Inauen, 2011). El estándar óptimo de esta relación de acuerdo al modelo del Ceaaces es 25:1.

Por otra parte, no siempre es deseable tener un profesor con el grado académico de doctorado o PhD para cubrir todas las demandas de la enseñanza; la participación de ayudantes de cátedra, doctorantes de la universidad o instructores que, dependiendo de cada universidad, pueden o no formar parte de la planta docente, es una modalidad común en la enseñanza universitaria. Esta diferenciación conduce a cierta demanda de horas de enseñanza anual por estudiante, que necesariamente tiene que empatarse con la capacidad de la universidad. Así, para cada disciplina académica y al menos para dos tipos de universidad –más orientadas a la investigación y más orientadas a la docencia- diferentes valores de la relación estudiantes /docentes serán necesarios para garantizar calidad en la enseñanza y el aprendizaje. Por consiguiente, si esta relación se va a usar para la formulación de políticas y asignación de recursos, esta tendrá que ser cuidadosamente analizada y revisada en función de la disciplina y su orientación específica (Schenker-Wicki & Inauen, 2011, pág. 48).

- *Eficiencia terminal*

Otro de los indicadores privilegiados de medición de la eficiencia de una universidad es la tasa de graduación de los estudiantes o eficiencia terminal, definida como el porcentaje de una promoción de estudiantes que se titulan en el ‘tiempo reglamentario’. Linn (2009) señala que la eficiencia terminal obedece a un sinnúmero de factores y añade que este indicador puede ser discriminatorio para ciertos grupos sociales en términos de género, estratos económicos y grupos étnicos minoritarios o estudiantes adultos. Diversas circunstancias (familiares, económicas y de origen) pueden alterar patrones normales de estudio. Por lo general, los estudiantes pertenecientes a estos grupos pueden requerir periodos más largos para completar sus estudios en la medida que ellos necesitan trabajar para suplementar sus ingresos, más esfuerzo para nivelarse o actualizarse en los conocimientos o simplemente dedicar tiempo a asuntos familiares o actividades domésticas. Entonces la observación de Morley (2004) en el sentido de que ciertos parámetros pueden tener diferentes valores para diferentes grupos es muy pertinente.

Por otra parte, este indicador puede penalizar a instituciones que sirven a una significativa fracción de estos grupos sociales y por consiguiente, estas pueden verse obligadas a desincentivar el enrolamiento de estudiantes pertenecientes a estos grupos. Más aun, este tipo de indicador puede crear presión en las universidades para bajar sus estándares de graduación y alcanzar en la evaluación un mejor nivel de desempeño en lo que se refiere sus tasas de graduación (Linn, 2009). Por último, es necesario señalar, como lo advierte Scott (2000), una contradicción entre indicadores de eficiencia de una universidad y el fenómeno de aprendizaje continuo que caracteriza la sociedad del conocimiento; lo que sugiere que el proceso educativo es permanente, por lo que carece de sentido cerrar un proceso abierto (*open-ended arrangement*) bajo criterios de efectividad⁶.

⁶ La eficiencia terminal óptima, de acuerdo al modelo del CEAACES, está en el rango del 40 por ciento y el 60 por ciento de una cohorte estudiantil. Una universidad que muestre valores inferiores al 40 por ciento y superiores al 60 por ciento es penalizada en la evaluación. Sin duda, un curioso razonamiento de evaluación de la eficiencia terminal de una institución de educación superior.

La conclusión que se desprende de las observaciones anteriores señala la necesidad de replantear una serie de ideas preconcebidas que realmente aportan poco a asegurar los objetivos de mejoramiento de la enseñanza superior. Como se anotaba anteriormente, la inclinación a favorecer cierto tipo de indicadores obedece más a la disponibilidad de información cuantitativa que puede ser fácilmente procesada y que otorga un sello de objetividad y precisión a la evaluación.

3. Insuficiencias metodológicas

De acuerdo al Proyecto de reglamento de los procesos de evaluación de carreras (CEAACES, Marzo, 2014), la acreditación y categorización de una carrera se sustenta en dos elementos básicos: (i) la evaluación del entorno de aprendizaje, y (ii) el Examen Nacional de Evaluación de Carreras – ENEC (art. 43). La estructura de evaluación se presenta en forma esquemática en el Grafico 1, a continuación.

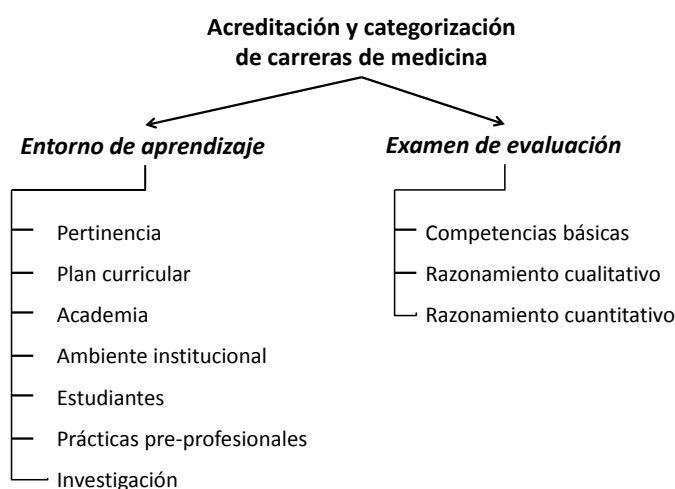


Grafico 1. Estructura del modelo de evaluación de carreras de medicina

El tratamiento analítico de una estructura de evaluación implica dilucidar una serie de dificultades conceptuales y metodológicas que, si bien pueden ser manejados con ayuda de herramientas apropiadas, estas no resuelven el problema de fondo de cualquier ejercicio de evaluación: ¿Cuáles son las características y atributos que se pretende evaluar y con qué objetivos? La evaluación no puede limitarse a un ejercicio de “ponderar y sumar”, sin un análisis de los supuestos que sustentan los arbitrajes sobre la compensación entre las variables del modelo, sobre los niveles de compromisos adoptados en la definición de pesos y ponderaciones y sobre las premisas que permiten acotar la ambigüedad de los atributos.

Estos elementos, que son en definitiva el andamiaje conceptual de un ejercicio de evaluación, cuestionablemente no figuran en el modelo de evaluación de carreras propuesto por el CEAACES (CEAACES, Agosto 2013) y todo parece indicar que tampoco han sido objeto de atención al momento de construcción del modelo de evaluación de las carreras de medicina (CEAACES, Junio, 2014). Esta omisión se explicaría, en parte, por un enfoque equivocado en el

tratamiento del problema, aquel de pretender buscar una “solución óptima” o la “mejor opción”, “de la manera más objetiva” (CEAACES, 2013, pág. 11), cuando precisamente, los conceptos multicriteriales constituyen una de las respuestas a la ortodoxia convencional de la racionalidad, y por lo tanto, del concepto de maximización de la utilidad.

Tres problemas conceptuales y metodológicos, inherentes a la ambigüedad de toda evaluación, exigen una explicación clara y fundamentada sobre los juicios y criterios adoptados por el ente evaluador en el diseño y aplicación del modelo. Estos son: niveles de compensación entre criterios, los compromisos y métodos adoptados en la definición de ponderaciones y el tratamiento a la indeterminación e incertidumbre de los elementos del problema en cuestión.

Compensación

La literatura sobre el análisis multicriterio distingue dos tipos de métodos: los métodos no-compensatorios y los compensatorios. Los primeros asumen que los criterios (o indicadores) son independientes entre sí de tal manera que cada uno contribuye de manera autónoma al objetivo de la evaluación. La contribución de cada criterio o indicador está determinada por su peso o ponderación. Este principio de independencia entre los criterios resulta, en muchas ocasiones, bastante restrictivo para la representación de las preferencias de los diferentes agentes involucrados en un proceso de evaluación. La introducción de la idea de compensación, es decir, el reconocimiento de interdependencias entre criterios (o indicadores) permite flexibilizar la rigidez inherente al principio de no-compensación entre los criterios de evaluación.

En el caso que nos ocupa, y a manera de ejemplo, el problema de compensación entre los criterios se plantea alrededor de las preguntas siguientes (Gráfico 2): ¿un desempeño excelente de una universidad con respecto a un indicador o criterio compensa su desempeño regular respecto a otros indicadores o criterios? ¿Qué se puede decir de esta universidad respecto a otra que presenta un desempeño aceptable en ambos casos?⁷

Una discusión sobre niveles de compensación entre los elementos de una estructura de evaluación va más allá de cuestiones formales que, en definitiva, tienen que ver con las ponderaciones que se asignan a los diferentes criterios e indicadores. Como se señaló anteriormente, el concepto de compensación tiene que ver, en parte, con la flexibilidad y, en definitiva, con la posibilidad de reconocer una cierta diversidad de los modelos universitarios. En efecto, un modelo de evaluación al delimitar una serie de criterios e indicadores, con sus respectivas ponderaciones y ‘funciones de valoración’, está implícitamente definiendo un modelo de universidad o carrera ‘ideal’ que sirve de referencia para determinar cuan ‘cerca’ o ‘lejos’ se encuentran los entes evaluados de este prototipo académico. La introducción de niveles de compensación rompe parcialmente la idea de un modelo único y rígido ya que el ‘score’ absoluto

⁷ El concepto de compensación en los métodos multicriterio, como se explica en el ejemplo, se refiere al grado de intensidad de la interrelación entre dos variables o criterios del modelo. Es diferente de la idea de compensación, introducida en el modelo del CEAACES bajo el nombre de *función de complejidad de la gestión* para modificar la valoración de ciertos indicadores en función del número de estudiantes. Este último artificio es altamente cuestionable tanto desde el punto de vista metodológico como conceptual.

deja de ser el parámetro predominante en la evaluación y es la idea de un comportamiento 'equilibrado' sobre un conjunto de criterios un factor que incide en los resultados.

| | Academia | Estudiantes |
|----------------------|----------|-------------|
| Universidad A | 85 | 20 |
| Universidad B | 55 | 40 |

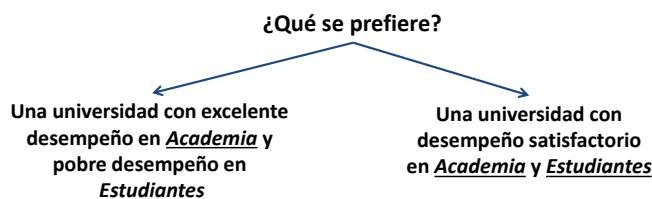


Grafico 2. El problema de compensación en el análisis multicriterial

En este sentido, cualquier modelo de evaluación debe incorporar elementos conceptuales y metodológicos que permitan un reconocimiento, al menos parcial, de la diversidad de las instituciones; es decir, distinguir universidades con misiones sociales, modalidades pedagógicas, campos disciplinarios, niveles educativos y hasta con concepciones de calidad diferenciadas. Esta heterogeneidad necesariamente conduce a reconocer objetivos y dinámicas diferenciadas y especializadas de los distintos tipos de universidades y por consiguiente, enfoques, criterios, parámetros y modelos de evaluación consistentes con esta pluralidad de instituciones (Villavicencio A. , 2012). El modelo aplicado en la reciente evaluación de universidades (CEAACES, 2013), por ejemplo, prescinde esta diversidad y plantea un modelo único, universal, aplicable indistintamente a cualquier circunstancia. De las cinco universidades ubicadas en la más alta categoría, dos corresponden un caso especial de instituciones de posgrado, con vocación internacional y que se rigen por un estatuto propio de acuerdo a convenios internacionales; otras dos universidades son instituciones con una vocación técnica, centradas en las ramas de la ingeniería y que, por lo tanto no pueden ser consideradas como un modelo de universidad (en el sentido amplio del concepto de universidad). Queda entonces una institución superior que el modelo de evaluación implícitamente la erige en el modelo universitario para las instituciones de educación superior del país.

Esta última observación de ninguna manera significa cuestionar un modelo universitario, muy legítimo, por cierto. El cuestionamiento va en el sentido de postular un modelo universitario como *el modelo* de la universidad ecuatoriana. El sistema de educación superior demanda urgentemente reconocer la coexistencia dentro del sistema de universidades con orígenes, roles y tradiciones diferentes. Este reconocimiento implica aprehender nuevas dimensiones de *calidad* que complejizan el ejercicio de evaluación y acentúan la dimensión diversa y multicausal de la pertinencia de la educación superior que no puede ser establecida, instrumentada ni fiscalizada por un enfoque único conceptual e institucional (Rama, 2008).

Indeterminación

Como fue anotado anteriormente, cualquier ejercicio de evaluación confronta el problema de vaguedad e indeterminación de los criterios. A esta ambigüedad conceptual se suma aquella relacionada con las escalas de valoración sobre la satisfacción de desempeño de una institución frente a los indicadores, ¿Cuál es el significado preciso de “cumplimiento total, parcial o deficiente” de una universidad respecto a un indicador X cuyo significado está rodeado de imprecisión y ambigüedad?

Por ejemplo, el modelo de evaluación define la carrera de medicina en los siguientes términos:

“Es el ámbito de la formación universitaria donde el estudiante adquiere los conocimientos, las habilidades y destrezas para ejercer la práctica médica con calidad y centrada en el paciente. El graduado de la carrera de Medicina es competente, efectivo y eficiente para tomar las mejores decisiones en el diagnóstico, manejo terapéutico, pronóstico y prevención de los principales problemas de salud que afectan a la sociedad ecuatoriana; de acuerdo con su nivel de competencia.” (CEAACES, Junio, 2014)⁸

Las preguntas que surgen de la lectura del enunciado anterior son obvias: ¿Cuáles son los parámetros que permiten calificar una práctica médica de calidad o afirmar que un graduado es competente, efectivo y eficiente? ¿Cómo discernir las mejores decisiones en el diagnóstico o determinar los niveles de competencia? Ante la falta de parámetros, criterios o referencias que permitan acotar, por lo menos en cierto grado, el lenguaje vago y confuso que caracteriza la mayor parte de los indicadores del modelo, el problema es resuelto por los evaluadores, profesionales convertidos en evaluadores universitarios. Se repite así un fenómeno ya observado en otras experiencias:

De un día a otro, académicos con trayectorias en su disciplina de origen (biólogos, arquitectos, sociólogos, médicos, etc.) se convirtieron en *evaluadores*. Bajo el lema “todos sabemos de evaluación, porque todos hemos sido evaluados alguna vez”, surgió una amplia generación de “expertos” en evaluación de programas, de los cuáles no se puede esperar más que la aplicación, en el mejor de los casos, de los conocimientos profesionales en los cuales han destacado con un sentido común en la evaluación (Díaz Barriga, 2006)

Lo que importa es calificar y, así como cualquier profesor califica a sus alumnos, cualquier profesor también puede evaluar. De ahí que cada evaluador tienda a realizar su ‘inspección’ bajo dictados del sentido común, siguiendo un patrón de instrucciones y “actuando mas como jurados que pares académicos” (Díaz Barriga, 2006). Entonces, son los evaluadores quienes califican, por ejemplo, los “niveles *adecuados* de interacción estudiantes-docentes”, quienes determinan las “cualificaciones *apropiadas* de los docentes”, como debe ser una “guía *adecuada* para la carrera” o, en definitiva, la “habilidad [de los docentes] para comunicarse, su entusiasmo para desarrollar programas más efectivos y su participación [adecuada?] en redes y sociedades profesionales”.⁹

⁸ Los subrayados son del autor

⁹ A manera de ejemplo, las citas corresponden a la definición del criterio Academia (CEAACES, Junio, 2014).

En este contexto, la precisión numérica¹⁰, el uso de fórmulas dudosas o de funciones arbitrarias son algunos de los elementos que pretenden suplir deficiencias, ambigüedades y subjetividades del análisis con una apariencia de rigor y objetividad. La focalización en este tipo de cuestiones puntuales significa reducir la evaluación a un ejercicio técnico que requiere la aplicación de la herramienta correcta. Esta tendencia es claramente resumida por Aboites, quien sostiene:

La evaluación permanece porque en apariencia ha resuelto el problema de la tecnología, ha logrado convencer que la búsqueda de la calidad no es una charlatanería, sino una ciencia y una tecnología sofisticadas que se expresan en instrumentos complicados, índices de confiabilidad, validaciones, formulas inescrutables, graficas, correlaciones y conclusiones, cuyo fundamento central –en un movimiento circular- son los mismos datos que produce. Se ostenta como científica, infalible y como epitome de la modernidad y el progreso. Es la tecnología de la diferencia nanométrica (Aboites, 2012, p. 913).

Sin embargo, esta aparente científicidad y neutralidad de la ‘metodología de evaluación’, presenta problemas engorrosos en el momento de categorizar o establecer un ‘ranking’ de las carreras evaluadas. En el modelo de evaluación propuesto, si, por ejemplo, se toman las dos dimensiones que determinan la acreditación de la carrera (examen de evaluación y entorno de aprendizaje) es perfectamente factible que dos carreras difieran escasamente en el desempeño obtenido respecto a estas dimensiones. Sin embargo, de acuerdo a la normativa del proceso (CEAACES, Marzo, 2014) de evaluación, los resultados, en términos de acreditación, son diametralmente opuestos (ver Grafico 3).

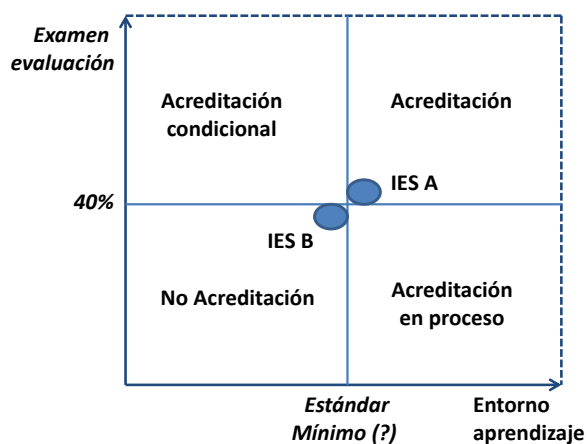


Grafico 3. El problema de proximidad en el análisis multicriterial

Pero no es únicamente la interpretación subjetiva de indicadores extremadamente ambiguos el factor que puede incidir en los resultados de la evaluación. La aplicación de herramientas, que en principio parecerían neutras, afectan los resultados dependiendo del

¹⁰ “Los pesos absolutos de los indicadores se calculan con una precisión de cuatro cifras decimales. Por efecto de presentación se muestran con una cifra decimal” (CEAACES, 2013)

algoritmo utilizado. Este es el caso, por ejemplo, del uso de los métodos de Análisis de Conglomerados, una técnica estadística, aplicada en la clasificación de las universidades en las distintas categorías (CEAACES, 2013). Varios algoritmos son posibles en la aplicación de estas herramientas y cada uno puede conducir a resultados diferentes. Así, la acreditación de una institución puede depender del tipo de algoritmo que se aplique. Siempre existe la posibilidad que el uso de una técnica tienda a exagerar o acortar las diferencias entre instituciones más allá de las evidencias estadísticas.

Ponderaciones

La asignación de las ponderaciones o pesos a los criterios e indicadores es, en definitiva, un ejercicio de dilucidación de preferencias que, en última instancia, responden a una concepción de un modelo de universidad. Se trata básicamente de un ejercicio de alcanzar compromisos entre una serie de objetivos parciales que no siempre apuntan en la misma dirección. De ahí la importancia, y condición básica de transparencia de un ejercicio de evaluación, explicitar con claridad los supuestos y consideraciones que sustentan las preferencias expresadas en las ponderaciones. Ahora bien, estas preferencias no siempre responden a los principios del ‘actor racional’ (estabilidad en el tiempo, transitividad, y perfectamente jerarquizadas), en el que se fundamenta la teoría de la utilidad que sustenta la metodología aplicada en la evaluación de las universidades (CEAACES, 2013) y el modelo genérico para la evaluación de carreras (CEAACES, Agosto 2013).

Las dos grandes limitaciones que enfrenta el comportamiento humano: la incertidumbre debida a la falta de información completa y los sesgos cognitivos por su capacidad de tratamiento limitada de la información, ponen en entredicho la racionalidad de sus preferencias (Harguindéguy, 2013). De ahí la necesidad de contar con instrumentos o herramientas de apoyo, que sin pretender resolver el problema de la racionalidad limitada, sirven como instrumentos de manejo de la complejidad al permitir transparentar la coherencia de los supuestos que están detrás de los ‘mágicos números’ de las ponderaciones.

Las herramientas y modelos son construcciones que a partir de un conjunto de hipótesis, generalmente restrictivas y parciales permiten abordar el problema bajo diferentes perspectivas. Esta limitación, inherente a cualquier modelo, puede ser superada mediante la aplicación de herramientas alternativas y/o complementarias que, en definitiva, permiten examinar la consistencia de hipótesis y, sobre todo, de resultados (el Cuadro 4 resume algunos de las metodologías de análisis aplicadas en el contexto del Mandato Constituyente No. 14). De ahí la necesidad de aplicar diferentes enfoques y métodos para evitar una suerte de determinismo metodológico. La evaluación de carreras no puede repetir el equívoco de la evaluación institucional de centrar la evaluación en la aplicación de un concepto único, sino analizar el problema de evaluación bajo enfoques complementarios y/o alternativos. Esta es otra condición para avalar la legitimidad e integridad del proceso de evaluación.

Cuadro 4. Conceptos y modelos aplicados bajo la evaluación del Mandato 14

| Enfoque | Objeto de análisis |
|---------------------------------|---|
| Teoría de la Utilidad | La aplicación de la teoría de la utilidad en un contexto de <i>racionalidad limitada</i> es el principio de aplicación de esta teoría a problemas de decisión que involucran objetivos múltiples. Los conceptos y herramientas del método permiten transparentar y manejar los supuestos subyacentes a los niveles de compensación entre los atributos u objetivos, así como los criterios de compromiso en la definición de sus preferencias (Keeney & Raiffa, 1993) |
| Lógica difusa | La lógica difusa es apropiada para el manejo de conceptos complejos y polimorfos que difícilmente pueden ser cuantificados y que se caracterizan por su ambigüedad o incertidumbre. El razonamiento con dichos conceptos no es claro y obvio, sino más bien difuso. El método se basa en la extensión de la lógica binaria a una lógica polivalente que provee los conceptos y herramientas para el manejo de conceptos y razonamientos ambiguos e imprecisos y permite el manejo de problemas caracterizados por diferentes grados de subjetividad (Chen & Hwang, 1992). |
| El Proceso Analítico Jerárquico | Procedimiento sistemático para el manejo de la complejidad de un problema mediante su estructuración en forma de una secuencia jerárquica. La importancia o preferencia de los elementos que constituyen la jerarquía (criterios, subcriterios, indicadores) es dilucidada sobre la base de <i>principios de discriminación y juicios comparativos</i> sintetizados en el método de la lógica de prioridades (Saaty & Kearns, 1985). |

4. Una evaluación centrada en ‘Evidencias’

Al igual que la evaluación anterior de las IES, la evaluación de la carrera tiene como eje central las ‘evidencias’ exigidas por el ente evaluador como prueba de cumplimiento y desempeño de la universidad respecto a los indicadores. La presentación de “actas”, “resoluciones”, “contratos”, “registros”, “certificados debidamente legalizados”, “memorandos”, “títulos refrendados” o “cualquier documento similar” son algunos de las constancias requeridas como prueba de la realización de un evento académico, de la presentación de una ponencia, del número de clases dictadas por los docentes o de una conferencia magistral dictada en la universidad, por ejemplo.

El volumen de ‘evidencias’ exigidas va más allá de la capacidad de procesamiento de una unidad administrativa de la carrera y, más aun, de la capacidad de análisis de un reducido grupo de evaluadores. Así, por ejemplo, la información requerida para valorar únicamente un indicador (ver Cuadro 5), de un total de cincuenta indicadores, requiere análisis socio-económicos sobre la incidencia de la carrera de medicina en el desarrollo local, regional y nacional, estudios de ‘mercado de trabajo’ y hasta estudios prospectivos sobre posibles problemas de salud que puedan presentarse en el futuro y las potenciales respuestas científico-tecnológicas para enfrentarlos. Todo este enorme esfuerzo para que un evaluador plasme en su informe que la “coherencia entre datos e informaciones y las políticas tomadas es total, parcial o bajo”.

Pero, el problema con las evidencias no se limita al enorme esfuerzo que deben dedicar los docentes y la administración en la preparación de la información exigida, sino que esta debe estar en formatos que pueda ser verificada “sin inconvenientes” por parte de los evaluadores. Así, la celeridad con la que el equipo evaluador procese las ‘evidencias’ es considerada como un indicador de la calidad académica de la carrera¹¹.

*Cuadro 5. Evidencias requeridas para la valoración del indicador
‘estado actual y prospectiva’ (CEAACES, Junio, 2014)*

- 1. Información general de la carrera:** Denominación de la titulación, modalidad, duración de los estudios, fecha de la primera aprobación de la carrera, denominación del ente aprobador, fecha de la última aprobación, nombre del ente o institución que la aprobó, tabla con los datos de los aspirantes a la carrera, estudiantes admitidos, número total de estudiantes, número de graduados durante los tres últimos años.
- 2. Análisis de pertinencia de la carrera.** Se debe mostrar, de qué manera el diseño de la carrera responde: a) Las necesidades del entorno provincial y local; b) Las necesidades del entorno nacional; c) Las necesidades del desarrollo científico-tecnológico; d) Los requerimientos de la planificación nacional y regional.
- 3. Resumen del estudio de la empleabilidad [sic] de los graduados:** universo de los graduados, tamaño de la muestra para el estudio, metodología(s) utilizada(s), tipo de trabajo desempeñado en relación con los años transcurridos desde la graduación, relación del trabajo u ocupación en la que se desempeña con los estudios realizados, ingreso mensual percibido, análisis de la situación laboral real de los graduados, etc.
- 4. En relación con la evolución prospectiva del área de la carrera:** Cuáles son los campos que se están desarrollando, que se prevé que se desarrollen a futuro, etc.; qué medidas se ha tomado con respecto al currículo.
- 5. Medidas que se ha tomado con respecto al currículo, con base en el análisis de la situación laboral real de los graduados.**

Toda esta política está creando un fenómeno nuevo en la academia ecuatoriana, el de las ‘evidencias’. En la evaluación realizada por el CEAACES muchas instituciones se vieron afectadas en la clasificación, ya sea por no disponer de una constancia que compruebe la realización de un evento académico, de un documento que demuestre las horas de clase dictadas por un docente, o el número “debidamente certificado” de bibliotecas virtuales, por ejemplo. Otras universidades se vieron también afectadas en su clasificación porque la constancia presentada por la universidad no

¹¹ El indicador “Información para la evaluación” (1.2.2) “evalúa la calidad de la información que la carrera proporciona al CEAACES para los procesos de evaluación y acreditación, en función de las variables del modelo que han podido ser verificadas sin inconvenientes, con base en la información y evidencias presentadas por la carrera”.

estaba en el formato exigido o era insuficiente para satisfacer el criterio de ‘cumplimiento’, de acuerdo a la escala de medición o valoración definida por la burocracia evaluadora.

Surge así en la academia ecuatoriana casi una nueva disciplina, a la que se podría denominar la “*evidenciología*”, definida como la ‘ciencia’ de recopilar, archivar y presentar evidencias académicas y en los formatos y contenidos exigidos por la agencia evaluadora. Entonces no resulta extraño que, en un ambiente de desconfianza hacia las universidades y bajo una evaluación de corte burocrático y controlador, estas se vean obligadas a montar estructuras administrativas para la gestión de las evidencias. Ahora resulta imprescindible presentar a la autoridad evaluadora, la constancia escrita, con sello, membrete y firma, que efectivamente certifique que el docente X ha presentado una ponencia, que en la universidad se ha dictado una conferencia magistral o que se ha impartido un seminario, etc. Sin la constancia respectiva, la ponencia o el seminario o la conferencia –no obstante el valor académico que tuvieran- no vale un solo ‘punto’. Así, la lógica de la agencia evaluadora se reduce a la fórmula:

$$\textit{evidencia} = \textit{calidad}.$$

Poco importa el contenido, pertinencia o trascendencia de los productos respaldados por las evidencias, lo que interesa es la producción bruta.

Por supuesto que la planta docente es la más afectada, porque los profesores, además de sus compromisos docentes y la creciente presión para publicar, deben dedicar cada vez más tiempo para elaborar tablas, llenar formatos puntualmente documentados y así, acumular ingentes cantidades de papel que respalden su desempeño y les permita estar preparados para responder a nuevas exigencias de una evaluación constante. Entonces, el volumen de productos funciona como un fetiche que hace innecesaria cualquier transformación que dé lugar a una mejor manera de hacer las cosas y, menos, generar más calidad en el quehacer institucional. Se está propiciando a que las instituciones funcionen para el llenado de formularios y tablas, lo que Porter denomina “universidades de papel” porque en los documentos son una cosa que no necesariamente ocurre en la realidad.

Quizá, más preocupante aun en este proceso de evaluación como control y vigilancia, es una pérdida gradual de las instituciones de educación superior de su capacidad para tomar decisiones mediante mecanismos establecidos y operados internamente, principalmente por sus cuerpos académicos y sin interferencia de intereses externos. Esta pérdida de capacidad se traduce en un proceso creciente de pérdida de autonomía del sistema universitario. Una evaluación organizada y administrada desde el Gobierno se instituye como el pivote referencia esencial para las acciones o reformas necesarias que debe emprender una institución para perfeccionar la docencia, aumentar su ‘eficiencia académica’, mejorar su organización e infraestructura. Los “planes de mejoras”, dictados y estrechamente fiscalizados por la institución evaluadora se convierten en el molde de la vida universitaria¹².

¹² La universidad es objeto de un meticuloso escrutinio del “Plan de desarrollo estratégico institucional”, “Plan operativo anual”, “plan de mejoras” y hasta de un “documento que contenga la planificación no ejecutada del POA” (ver indicador “Estado actual y prospectiva: descripción de las evidencias).

La universidad asiste impasiblemente a un proceso en el que pasa de ser una institución relativamente autónoma a una mera dependencia de la burocracia gubernamental, sujeta a la supervisión de la agencia evaluadora; proceso que viene a ser un símbolo de la enorme burocratización y subordinación que la desvanece como factor decisivo en la vida nacional. De hecho, se crean las bases de un nuevo esquema corporativo entre el Estado y las instituciones de educación superior, estas últimas como productoras de servicios y conocimiento y el primero como regulador de la competitividad y el mercado. En estas circunstancias, la universidad, como acertadamente lo señala Aboites (2012, pág. 741), “es objeto de un curioso procedimiento ‘quirúrgico’ que le quita toda substancia, una especie de lobotomía que la reduce a un estado deplorable”.

5. Conclusiones

La discusión anterior plantea una serie de cuestionamientos, tanto de carácter conceptual como metodológico, al modelo genérico de evaluación de carreras (CEAACES, Agosto 2013) y, en particular, al modelo de evaluación de la carrera de medicina (CEAACES, Junio, 2014).

- a) Existe una evidente redundancia entre los indicadores. Varios indicadores se refieren a un mismo atributo o característica objeto de evaluación, situación que introduce un sesgo en la integridad del ejercicio evaluatorio. Al no observarse el principio de exclusividad en los elementos de la estructura de evaluación, algunas universidades pueden resultar favorecidas, mientras otras pueden verse afectadas negativamente.
- b) Varios indicadores evalúan características y atributos que se encuentran claramente especificados en normas y reglamentos. Por consiguiente, carece de sentido valorar el cumplimiento de una norma en términos de cumplimiento total, parcial o insatisfactorio. En este caso, cualquier ‘función de utilidad’ o escala de tipo cualitativo resulta superflua.
- c) Varios indicadores (estudiantes y docencia, eficiencia terminal, investigación, publicaciones, por ejemplo), tal como están definidos responden a un conjunto de hipótesis no comprobadas, simplificadoras de la realidad universitaria y de impacto mínimo o irrelevante sobre la enseñanza (Terenzini & Pascarella, 1994). Otros indicadores resultan cuestionables como ‘eficiencia terminal’ (Linn, 2009) o simplemente no tienen nada que ver con los objetivos de la evaluación como es el caso del indicador ‘información para la evaluación’¹³.
- d) Un modelo de evaluación debe ser capaz de aprehender enfoques académicos heterogéneos, esquemas pedagógicos diversos y hasta concepciones de calidad diferenciadas (Villavicencio A. , 2013). El modelo del CEAACES, sin embargo, reivindica un modelo único de carrera como el arquetipo de carrera para las universidades del país. Esto constituye una arbitrariedad que requiere ser corregida. La introducción en el modelo de criterios de

¹³ Un análisis sobre la pertinencia y relevancia de los indicadores específicos a la carrera de medicina, en particular los referentes a los criterios de ‘prácticas pre-profesionales’ y ‘plan curricular’ está fuera del alcance del presente trabajo. Corresponde a las facultades de medicina pronunciarse al respecto.

compensación entre los elementos de la estructura de evaluación permitiría corregir, por lo menos parcialmente, esta peligrosa limitación.

- e) Imprecisiones y ambigüedades del lenguaje son inevitables. Estas, sumadas a sesgos cognitivos, crean amplios márgenes de subjetividad al momento de formular preferencias. Estos márgenes pueden ser acotados si los criterios e indicadores son definidos en forma operacional (Keeney & Raiffa, 1993); es decir, de tal manera que su significado permita i) demostrar su pertinencia y relevancia; ii) deducir con claridad las implicaciones de los criterios de evaluación respecto a los objetos evaluados; y iii) expresar con claridad y transparencia su contenido y alcance con el fin de facilitar el dialogo. Los criterios e indicadores del modelo propuesto por el CEAACES se caracterizan por una enorme ambigüedad en su formulación. Casi todos ellos están definidos en función de términos como “adecuados”, “coherentes”, “calidad”, dejando espacios prácticamente ilimitados para una interpretación discrecional por parte de los evaluadores. La reformulación de los criterios e indicadores en términos operacionales para los objetivos de la evaluación resulta indispensable para asegurar un mínimo de legitimidad del proceso.
- f) El empleo de herramientas y técnicas multicriteriales constituye otro mecanismo para mitigar las imprecisiones y ambigüedades; pero no en el sentido de ‘cajas negras’ que las eliminan, sino como herramientas de manejo de la complejidad que permiten transparentar la lógica de los razonamientos y preferencias subyacentes en la evaluación. Sin embargo, los conceptos y técnicas no son neutras ni universales. Estas parten de supuestos restrictivos que privilegian aspectos parciales del problema. De ahí la necesidad de aplicar diferentes enfoques y métodos para evitar una suerte de determinismo metodológico. La evaluación de carreras no puede repetir el equívoco de centrar la evaluación en la aplicación de un concepto único, sino analizar el problema de evaluación bajo enfoques complementarios y/o alternativos. Esta es otra condición para avalar la legitimidad e integridad del proceso de evaluación.
- g) La observación anterior conduce a la necesidad de introducir un análisis de sensibilidad como elemento imprescindible del ejercicio de evaluación. Aun en el caso de información cuantitativa que sustenta ciertos indicadores, esta se encuentra rodeada de incertidumbre que exige, en algunos casos, un tratamiento probabilístico. Analizar detalladamente la sensibilidad de los resultados de una evaluación frente a rangos de incertidumbre resulta indispensable.

Pero, además de los aspectos formales, el debate sobre la evaluación universitaria tiene que ver con el método o enfoque de evaluación al que se hizo referencia anteriormente (ver sección 2). El objetivo de mejorar el sistema de educación superior requiere una aproximación a las universidades no como entidades fijas caracterizadas por un conjunto de atributos, sino como entidades que se definen ellas mismas de diferentes maneras y modifican su identidad y preferencias como resultado de experiencias y aprendizajes (Abbott, 2001). La agencia evaluadora

debe comprender que el sistema universitario, desde la evaluación bajo el Mandato Constituyente 14, ha experimentado radicales cambios que demandan nuevos enfoques, conceptos y métodos para lograr que las reformas introducidas empiecen a producir los cambios esperados. Independientemente de la pertinencia de las normas en vigencia, el sistema universitario cuenta con un conjunto de reglamentos, cuya observación no puede ser objeto de apreciaciones subjetivas sobre su *cumplimiento total, parcial o deficiente*. Esta confusión de *meritos* con cumplimiento de normas tiene sus raíces en espíritu de vigilancia y control de los reglamentos-instructivos, que al pretender normar todos los aspectos de la vida académica, cierran los espacios para imaginar y emprender programas de mejoramiento gradual de la enseñanza universitaria.

Entonces, en lugar de una evaluación, como un ejercicio técnico y neutro de control y fiscalización, el mejoramiento del sistema universitario debe partir de definiciones flexibles y amplias de calidad, que den cuenta de la diversidad de instituciones, de estudiantes y demandas. Esto implica el análisis de nuevos modelos institucionales que permitan avanzar y especificar esas definiciones de calidad como sinónimo de pertenencia; la necesidad de vincular el proceso con otros instrumentos de política, pero no de manera lineal o directa, sino de manera complementaria con mecanismos de apoyo, regulatorios, de financiamiento u otros, pero siempre tomando en consideración las implicaciones de dicha vinculación (Pires & Lemaitre, 2008). Si no se cambia de perspectiva, la evaluación, cualquiera que sea su objetivo, corre el riesgo de perder su legitimidad.

Referencias

- Abbott, A. (2001). *Time Matters: On Theory and Method*. Chicago: The University of Chicago Press.
- Abaites, H. (2012). *La medida de una nación: los primeros años de la evaluación en México*. México: Universidad Autónoma Metropolitana.
- Astin, A. (1996). Involvement in learning revisited. *Journal of College Student Development*, 40 (5), 587-597.
- CEAACES. (2013). *Informe General sobre la Evaluación, Acreditación y Categorización de las Universidades y Escuelas Politécnicas*. Quito: Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior.
- CEAACES. (Junio, 2014). *Modelo para la Evaluación de las Carreras de Medicina - versión matricial*. Quito.
- CEAACES. (Agosto 2013). *Modelo para la Evaluación de las Carreras Presenciales y Semi-presenciales de las Universidades y Escuelas Politécnicas del Ecuador*.
- CEAACES. (Marzo, 2014). *Proyecto de Reglamento de Evaluación y Acreditación de Carreras*.
- Chen, S.-J., & Hwang, C.-L. (1992). *Fuzzy Multiple Attribute Decision making*. Berlin: Springer-Verlag.
- Dias Sobrinho, J. (2008 b). Calidad, Pertenencia y Responsabilidad Social de la Universidad Latinoamericana y Caribena. In A. L. Gazzola, & A. Didriksson (Eds.), *Tendencias de la Educación Superior en América Latina y el Caribe* (pp. 87-112). Caracas: IESALC-UNESCO.
- Díaz Barriga, A. (2006). Evaluación de la educación superior. Entre la compulsividad y el conformismo. In L. Todd, & V. Arredondo (Eds.), *La Educación que México necesita: visión de expertos*. Monterrey: Centro de Altos Estudios e Investigación Pedagógica.
- Dill, D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education*, 49, 495-533.

- Donoghue, F. (2008). *The Last Professors: the corporate university and the fate of humanities*. New York: Fordham University Press.
- Grin, J., Rotmans, J., & Schot, J. (2010). *Transitions to Sustainable Development, New Directions in the Study of Long Term Transformative Change*. New York: Routledge.
- Harguindéguy, J.-B. (2013). *Análisis de Políticas Públicas*. Madrid: Tecnos.
- Keeney, R., & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York: Cambridge University Press.
- Linn, D. (2009). *Measuring student achievement at postsecondary institutions*. Issue Brief, NGA Center for BEST PRACTICES, Education Division, Washington D.C.
- Morley, L. (2004). *Quality and Power in Higher Education*. Philadelphia: Open University Press and McGraw Hill Education.
- Pires, S., & Lemaitre, M. J. (2008). Sistemas de Acreditación y Evaluación de la Educación Superior en América latina y el Caribe. In A. L. Gazzola, & A. Didriksson (Eds.), *Tendencias de la Educación Superior en América Latina y el Caribe* (pp. 297-318). Caracas: IESALC/UNESCO.
- Rama, C. (2008). *Tendencias de la Educación Superior en América Latina y el Caribe en el Siglo XXI*. Lima: Asamblea Nacional de Rectores.
- Saaty, T., & Kearns, K. (1985). *Analytical Planning: The organization of systems*. Pittsburg, PA: RWS Publications.
- Schenker-Wicki, A., & Inauen, M. (2011). The economics of teaching: what lies behind student-faculty ratios? *Higher Education Management and Policy*, 23 (3), 31-50.
- Scott, P. (2000). Globalisation and higher education: challenges for the 21st century. *Journal of International Studies in International Education*, 4 (1), 3-10.
- Terenzini, P., & Pascarella, E. (1994). Living with myths: Undergraduate education in America. *Change*, 26 (1), 28-32.
- Villavicencio, A. (2013). ¿Hacia donde va el proyecto universitario de la Revolución Ciudadana? In J. Cuví, D. Machado, A. Oviedo, & N. Sierra (Eds.), *El correísmo al desnudo* (pp. 216-231). Quito: Montecristi Vive.
- Villavicencio, A. (2012). *Evaluación y acreditación en tiempos de cambio: la política pública universitaria en cuestionamiento*. Quito: Cuaderno de Trabajo. Instituto de Altos Estudios Nacionales.