

El *homo academicus* y la máquina de sumar: profesores universitarios y la evaluación cuantitativa del mérito académico

HÉCTOR VERA*

Este artículo analiza el proceso de cuantificación en la evaluación del mérito académico, tal como se presenta en el caso de la Universidad Autónoma Metropolitana (UAM), una institución pionera a nivel mundial en la aplicación de métodos cuantitativos de evaluación de profesores-investigadores. Se realizó un análisis del funcionamiento del tabulador para ingreso y promoción de la UAM, se condujeron 15 entrevistas a profundidad y se llevó a cabo una encuesta de respuesta voluntaria entre académicos vía Internet. El artículo detalla las transformaciones en los instrumentos de evaluación y los cambios de comportamiento que éstos han inducido entre los profesores. Se describe también cómo las evaluaciones cuantitativas le permiten a la institución legitimar sus procesos de evaluación y cómo facilita la comparabilidad entre los sujetos evaluados; al mismo tiempo, la evaluación cuantitativa permite a los académicos conmensurar los distintos productos de su trabajo y les da certeza sobre los posibles resultados de la evaluación.

This article analyzes the process of quantification of academic merit. The case of the Metropolitan Autonomous University (Universidad Autónoma Metropolitana), which was a global forerunner in the use of quantitative evaluation methods for professors, is examined in depth. Data was obtained by examining the listing tables employed by the university to quantify academic products, and by semi-structured interviews and a voluntary survey among full time professors. The article details the changes in the instruments of evaluation used by the university and the impact that those have had on faculty performance. It also describes how quantitative evaluations allow university administrators to legitimate the evaluations and facilitate the comparability of the evaluated subjects; on the other hand, quantitative evaluations let faculty commensurate their different academic products and give them certainty about the possible outcomes of the evaluation process.

Palabras clave

Evaluación cuantitativa
Universidades
Universidad Autónoma
Metropolitana
Profesores
Administración
universitaria

Keywords

Quantitative evaluation
Universities
Metropolitan
Autonomous University
Professors
University administration

Recepción: 15 de marzo de 2016 | Aceptación: 28 de junio de 2016

* Investigador del Instituto de Investigaciones sobre la Universidad y la Educación de la Universidad Nacional Autónoma de México (UNAM). Doctor en Sociología y Estudios Históricos por la New School for Social Research. Líneas de investigación: usos sociales de los sistemas de medición, evaluación cuantitativa del mérito académico. Publicaciones recientes: (2016), "Rebuilding a Classic: *The Social Construction of Reality* at 50", *Cultural Sociology*, vol. 10, núm. 1. DOI: <http://dx.doi.org/10.1177/1749975515617489>; (2016), "Weights and Measures", en B. Lightman (coord.), *Blackwell Companion to the History of Science*, Oxford, Blackwell Publishers. CE: hectorvera@unam.mx.

INTRODUCCIÓN¹

Para la administración de instituciones de educación superior se ha vuelto crucial encontrar métodos operativos —y aceptables— para determinar qué es “excelencia”, “calidad”, “mérito” y “productividad” entre sus profesores e investigadores. Esto es importante, entre otras cosas, para distribuir de manera clara y poco controversial los recursos y reconocimientos disponibles en cada institución: contrataciones, promociones, becas, primas al desempeño, estímulos y premios. Durante los últimos 30 años, las instituciones de educación superior han optado de manera creciente por utilizar *indicadores cuantitativos* que intentan *medir y dar una expresión numérica simple e inequívoca* a cuan sobresalientes (o insuficientes) son los logros de los docentes e investigadores. Como resultado, se han diseñado y aplicado numerosas “tecnologías del mérito” (Carson, 2004), como tabuladores, índices de impacto, factores numéricos, *rankings* e indicadores bibliométricos, que se han convertido progresivamente en herramientas cotidianas —y en ocasiones en guías prioritarias o exclusivas— de la evaluación del quehacer académico.

Esta tendencia ha implicado que a los elementos tradicionales con los que se valora el mérito de la labor y trayectoria académicas (como títulos, condecoraciones, reputación de las revistas y casas editoriales donde se publican libros y artículos, prestigio de los asesores y universidades que otorgan los grados académicos, etcétera) se han sumado instrumentos primariamente *cuantitativos* que cobran cada vez mayor peso en las decisiones institucionales.

La cantidad de actores que están involucrados en estos procesos es muy amplia (con variantes según el país) e incluye a los académicos individuales, dictaminadores,

administradores universitarios, directores de presupuesto, consejos nacionales de ciencia y agencias nacionales e internacionales de financiamiento (Cronin y Sugimoto, 2014). Las disputas entre estos actores sobre la utilidad (o ineficacia) de indicadores cuantitativos para comprender y valorar la actividad académica y científica se han multiplicado en los últimos años a nivel internacional (Abbott *et al.*, 2010; Van Noorden, 2010): mientras algunos abogan por continuar perfeccionando los métodos cuantitativos de evaluación, con la idea de garantizar por ese medio evaluaciones menos arbitrarias y superficiales (Ricker, 2015), otros advierten que este tipo de mediciones es incompleto y que su puesta en práctica irreflexiva ha tenido consecuencias no buscadas e indeseables (Hicks y Wouters, 2015; Wilsdon, 2015).

Pierre Bourdieu le dio el sugestivo título de *Homo academicus* a su estudio sobre la estructura y funcionamiento del mundo universitario. Ahí analizó, entre otras cosas, las especies de poder que son eficaces en ese campo y las trayectorias de los agentes que se encuentran en él. Para ello describió las luchas de los académicos por controlar tanto los recursos que permiten el entrenamiento y las promociones dentro de la universidad, como los medios para la investigación y para obtener prestigio (Bourdieu, 2008; Swartz, 1997). En el medio universitario mexicano, enfrentarse a las evaluaciones cuantitativas y a sus respectivas “máquinas de sumar” (*i.e.*, los instrumentos con los que se le otorga un valor numérico a las actividades académicas) son un requerimiento inescapable para que los agentes logren acumular parte del capital (*e.g.*, becas, plazas, asensos, premios, posiciones honoríficas) que utilizan para tener una posición en el campo y para luchar por más recursos. En otras palabras, las evaluaciones cuantitativas están en el corazón de las trayectorias y conflictos del “*homo academicus*”.

¹ Deseo agradecer el valioso apoyo que recibí para esta investigación de parte de Lorena Cruz, Dámaris Argumedo y Mariana Nieves Uribe. Este trabajo tuvo su primera versión en la ponencia “La medición cuantitativa del mérito. El caso de la Universidad Autónoma Metropolitana” (Vera, 2015).

La forma en que se ha aceptado, negociado o resistido la penetración de las mediciones del mérito no han sido consistentes ni coherentes entre sí. Por un lado, este tipo de evaluaciones han sido admitidas y promovidas por una franja de los mismos académicos (con el apoyo o presión de cuerpos administrativos dentro y fuera de las universidades). Al mismo tiempo, las evaluaciones han sido objeto de críticas y objeciones recurrentes.² En particular, la evaluación del mérito académico por medio de tabuladores, como el que se utiliza en la UAM, ha sido criticada por los académicos porque muchas veces no se considera dentro de los tabuladores actividades importantes que los profesores-investigadores realizan en pro de los estudiantes o de la universidad. Aunque las universidades enfatizan que sus labores sustantivas son tres (docencia, difusión de la cultura e investigación) se premia de manera desproporcionada a la investigación (esto sumado al problema de que los tabuladores no siempre se actualizan y no consideran actividades que sólo recientemente cobraron relevancia, como los posdoctorados, por ejemplo). También se ha señalado que las evaluaciones cuantitativas ligadas a un esquema de estímulos promueven simulación, corrupción académica, individuación de la producción científica y el abandono de los proyectos de investigación a largo plazo en favor de investigaciones que pueden dar resultados en poco tiempo (Ordorika, 2004). A nivel internacional se ha señalado que la evaluación cuantitativa sustituye, en vez de apoyar, la valoración cualitativa de los expertos, y que la evaluación queda dirigida por los datos en lugar de por el juicio (Hicks y Wouters, 2015). Sin embargo, estos cuestionamientos no han articulado una visión suficientemente robusta que pueda menguar las prácticas y fuentes de legitimidad de las mediciones cuantitativas.

El objetivo de este artículo es contribuir a una evaluación crítica de los efectos de las evaluaciones cuantitativas (que son aceptadas irreflexivamente por unos y demonizadas precipitadamente por otros), lo que incluye preguntarse por las consecuencias de su uso en la práctica de individuos e instituciones, lo mismo que interrogar por qué ese particular método de evaluación ha sido aceptado por los agentes involucrados en el universo de las instituciones de educación superior. El artículo explica el proceso de cuantificación en la evaluación del mérito, tal como se ha presentado en la Universidad Autónoma Metropolitana (UAM). Para esto analiza la instauración y las transformaciones en sus instrumentos de evaluación, y las acciones y los cambios de comportamiento que éstos han inducido entre investigadores y administrativos. Este análisis intenta ayudar a explicar qué han hecho las instituciones y los académicos para adaptarse, negociar o rechazar el carácter numérico de los nuevos tipos de evaluación.

REPENSAR LA CUANTIFICACIÓN

En los debates recientes sobre las políticas de evaluación se ha prestado atención insuficiente al *carácter específicamente cuantitativo* que las evaluaciones han adquirido en últimas décadas, por lo que se requieren estudios detallados para sopesar cuáles han sido las consecuencias de la evaluación cuantitativa.

El acenso de las evaluaciones cuantitativas se conjuga con el actual contexto político en el que han nacido controversias por la imposición y generalización de métodos de evaluación universal de los maestros de educación básica que están guiados por principios de valoración uniforme y estandarizada. Este artículo pretende agregar una nueva dimensión a los estudios que se están realizando en centros

² Se han cuestionado los *rankings* universitarios (Lloyd, 2011; Ordorika y Rodríguez, 2010; Enders, 2015), las compensaciones condicionadas al mérito (Galaz Fontes y Gil Antón, 2013), la evaluación en la educación superior (Díaz-Barriga, 2008) y los exámenes estandarizados (Aboites, 2012).

dedicados a la investigación educativa sobre esa filosofía de *normalización*³ de la evaluación.

Por otra parte, entender el aspecto cuantitativo de los procesos de evaluación es importante analíticamente en dos niveles distintos. Primero, socialmente: los números tienen una vida social particular en términos de su prestigio (bien ganado o no) como garantes de “objetividad”, “exactitud”, “confianza”, “racionalidad” y “universalidad” (Porter, 1996); con ello, la demanda para que se cuantifiquen los fenómenos sociales, como la educación, se ha incrementado (Vera, 2016). Sin embargo, los científicos sociales han prestado poca atención al origen, impacto y propagación de la cuantificación y de los nuevos “regímenes de medición” (Vera, 2005).

Segundo, cognitivamente: aprehender algo cuantitativamente implica una serie de procesos cognitivos que difieren significativamente de otros tipos de conocimiento (cualitativo, intuitivo, etcétera). Medir cuantitativamente el mérito implica producir diferencias cuantitativas entre realidades que son relativamente homogéneas en sus cualidades; y a la hora de hacer evaluaciones se utilizan esas diferencias cuantitativas producidas por la medición para demarcar y segregar entidades (*i.e.* clasificar y jerarquizar). Con esto, la producción universitaria está siendo incluida en una visión del mundo en la que lo “real” se ve como sinónimo de lo “medible” (Espeland y Stevens, 2008), al mismo tiempo que se emplean las mediciones para instrumentar y justificar la repartición jerarquizada de bienes simbólicos y económicos.

Cuando hablamos de “evaluación cuantitativa del mérito académico” hace falta clarificar los elementos clave de esa expresión. En primer lugar está la “evaluación cuantitativa”, que refiere a la estimación y comunicación de un valor por medio de números —lo que presupone la creencia de que aquello que se está

evaluando puede ser aprendido por medio de una cantidad (*i.e.*, a través de una propiedad que es susceptible de ser contada o medida)—. En segundo lugar, está el “mérito académico”. *Mérito* es una palabra volátil, pero muy significativa en el medio universitario; no se refiere únicamente a su significado primario de actividades dignas de reconocimiento y por las que se merece obtener algo; está también presente en expresiones que designan figuras y posiciones en el mundo académico, como la del profesor o investigador *emérito* (aquel que merece disfrutar de una recompensa debida a los méritos demostrados durante su tiempo de servicio); o en el término *meritorio*, que nombra cierto tipo de trabajo no remunerado en las universidades, como las “adjuntías meritorias” (es decir, ayudantías sin sueldo que realizan jóvenes aspirantes a ingresar al mundo académico con el fin de adquirir habilidades y “hacer méritos”). Así, la “evaluación cuantitativa del mérito académico” implica identificar cuáles son las actividades académicas dignas de reconocimiento y *calcular* cuánto valor merecen. Ese cálculo se realiza valuando diferentes actividades y productos con un estándar de medición común —lo que permite traducir en cantidades o magnitudes las diferencias o similitudes entre los elementos evaluados— (Espeland y Stevens, 2008).

METODOLOGÍA

Parte de los datos empíricos utilizados en esta investigación se obtuvieron a través de la realización de 15 entrevistas semiestructuradas con profesores-investigadores de tiempo completo de una de las cinco unidades académicas de la UAM. Las entrevistas se realizaron entre académicos de la División de Ciencias Básicas e Ingeniería y de la División de Ciencias Sociales y Humanidades.⁴ Sólo uno de los investigadores entrevistados forma

³ “Normalización” en el sentido de regularizar y estabilizar por medio de criterios invariables.

⁴ Se menciona frecuentemente que los modos de evaluación como el del tabulador de la UAM favorecen el modo de producción académica que es afín a las así llamadas “ciencias duras”, pero que es impropio para otras áreas del

parte de otra unidad académica de la UAM, pero fue consultado por haber presidido una comisión dictaminadora de área que evalúa a los profesores de todas las unidades de la UAM pertenecientes a esa área específica. Los nombres de los entrevistados y de la unidad académica donde laboran los académicos se han mantenido en el anonimato.

Una segunda fase de recolección de datos se realizó a través de una encuesta de respuesta voluntaria entre profesores-investigadores que dan clases en cuatro licenciaturas, tanto de la División de Ciencias Básicas e Ingeniería como de la División de Ciencias Sociales y Humanidades. La encuesta consistía en un cuestionario de 15 preguntas que podía ser respondido por Internet. Se enviaron invitaciones por correo electrónico para responder el cuestionario a 235 académicos. El porcentaje de respuesta fue del 10.2 por ciento.⁵

Para indicar la procedencia de las opiniones vertidas en las entrevistas y en los cuestionarios que son citados textualmente en el artículo, se utiliza la letra *E* para las entrevistas y la letra *C* para los cuestionarios, junto con un número que indica el orden cronológico en que las entrevistas fueron realizadas y en que los cuestionarios fueron recibidos (así, por ejemplo, una cita de la segunda entrevista realizada es identificada como E2).

Finalmente, se llevó a cabo una comparación entre el tabulador para ingreso y promoción del personal académico (TIPPA) que está

actualmente en uso para evaluar a todos los profesores de tiempo completo de la UAM, y el tabulador “cualitativo” anterior, que fue sustituido en 1985 (UAM, 2013). Adicionalmente, se analizaron los cuadernillos con los “Criterios de dictaminación” que se emplean en la comisión dictaminadora de ciencias sociales y en la de ciencias básicas (UAM, 2003; UAM, 2005), que sirven para complementar los criterios del TIPPA.

LA UNIVERSIDAD AUTÓNOMA METROPOLITANA Y LA EVALUACIÓN CUANTITATIVA DEL MÉRITO

La UAM fue una de las instituciones pioneras en la aplicación de métodos cuantitativos para la evaluación de profesores investigadores.⁶ En esta Universidad se realizan varias evaluaciones para ingreso, promoción y obtención de becas siguiendo un tabulador cuantitativo que sirve para todos los académicos, para todas las disciplinas y para las cinco unidades académicas. La introducción de este método cuantitativo de evaluación se realizó en 1985, en medio de una transformación general de los métodos de contratación y promoción en la UAM (Ibarra y Rondero, 2005; López Zárate *et al.*, 2000). La introducción de las evaluaciones primordialmente cuantitativas en la UAM se inició sólo un año después de la creación del Sistema Nacional de Investigadores del CONACyT (que apareció en 1984) y antecedió

conocimiento. Con esto en mente, se decidió seleccionar a profesores de dos distintas divisiones de la UAM (la de Ciencias Básicas e Ingeniería y la de Ciencias Sociales y Humanidades) para observar si aparecían discrepancias en cómo se utilizaba y cómo se percibía el instrumento de evaluación en distintas áreas disciplinares.

5 Se entrevistó y se le envió el cuestionario de respuesta voluntaria a profesores-investigadores de la UAM con nombramiento de titular por tiempo indeterminado (es decir, según la terminología de otras instituciones, académicos numerarios, con definitividad o con *tenure*). Entre los entrevistados se encontraba personal que ocupa o ha ocupado los siguientes puestos administrativos y directivos dentro de la universidad: jefatura de área, jefatura de departamento, coordinación de licenciatura, coordinación de posgrado, representante de departamento en el consejo académico, miembro de comisión dictaminadora, secretaria de unidad y rectoría de unidad. La antigüedad en la universidad de estos académicos es, en promedio, de 23 años.

6 Los sistemas de evaluación de los académicos universitarios en México comenzaron a tomar su forma actual tras una serie de profundas transformaciones iniciadas desde fines de los años setenta y consolidadas en la década siguiente. Por un lado, se realizaron reformas legislativas para reducir la influencia de los sindicatos en los procesos de contratación y promoción de los profesores e investigadores universitarios; por otra parte, ante la crisis económica nacional se diseñaron mecanismos de recuperación salarial para hacerle frente al deterioro de los ingresos de los académicos y para evitar la así llamada “fuga de cerebros” (Ordorika, 2004). Muchos de estos mecanismos tuvieron la forma de “pago por mérito”, con los que se otorgaba una gratificación (adicional al sueldo base) a aquellos académicos que librarán exitosamente procesos de evaluación *ad hoc*.

por tres años al inicio de la deshomologación salarial en la UAM (con la cual se otorgaba una compensación económica adicional a algunos académicos para reconocer su desempeño prominente y diferenciar su trabajo) (Rondero, 2007).

En general México fue pionero en la implementación de las políticas de pago por mérito en las universidades, tanto por su extensión dentro del sistema universitario como por el impacto económico que éstas han tenido en los ingresos de los académicos (Ordorika, 2004). La introducción del TIPPA en la UAM formó parte de ese cambio estructural; fue la primera universidad en México en poner en práctica estas políticas. A nivel internacional la UAM fue igualmente pionera en estos rubros, particularmente en lo que se refiere a la cuantificación de la evaluación del trabajo académico. Tan solo como punto de referencia, se puede apuntar que la reforma de la UAM antecedió en un año a la creación del Research Assessment Exercise,⁷ que se utiliza desde 1986 para evaluar la calidad de las investigaciones científicas que se realizan en las instituciones de educación superior en Inglaterra y es uno de los programas más visibles y discutidos internacionalmente.⁸ Igualmente, la reforma de la UAM llegó casi tres décadas antes de que el Departamento de Educación de Estados Unidos lanzara el *College Ratings Framework*, otra iniciativa de cuantificación educativa que ha atraído ampliamente los reflectores.

El método de evaluación de la UAM, a través del uso del tabulador, sirve no sólo como un “recuento curricular de méritos” (García Salord, 2013), sino como una *cuantificación*

curricular de méritos, donde los profesores de tiempo completo (o los candidatos a una plaza) ordenan los resultados de su trabajo (publicaciones, clases, conferencias, presentaciones de libros, dirección de tesis, asesorías de servicio social, coordinación de congresos, arbitraje de artículos, participación en órganos colegiados y un largo etcétera) llenando las casillas de un tabulador que asigna un puntaje (o, en ciertos casos, un rango de puntos) predeterminado e inmodificable para cada actividad.

Con la mencionada reforma de 1985, en la UAM se decidió sustituir el tabulador hasta entonces en uso (Tabla 1) por un segundo tabulador. El primer tabulador no era propiamente cuantitativo, y sólo detallaba la escolaridad y la experiencia docente y de investigación que hacían falta para aspirar a cada categoría y nivel académico. Por ejemplo, un profesor “asociado C” debía tener título de licenciatura, 50 por ciento de los créditos de la maestría, cinco años de experiencia docente, dos programas de clase y tres artículos publicados.

La reforma culminó en el actual tabulador para ingreso y promoción del personal académico (TIPPA). Una de las características más visibles de este nuevo tabulador fue su carácter cuantitativo; por ejemplo, en este nuevo tabulador la escala para la fijación de categorías se indica siempre como una cantidad de puntos. Así, los requisitos para “profesor asociado C” quedaban ahora definidos como “19,100 puntos”.

El TIPPA sirve entonces para fijar la categoría y el nivel del personal académico de acuerdo a una escala de puntos (Tabla 2).⁹ La

⁷ El RAE fue remplazado en 2008 por el Research Excellence Framework, que cumple esencialmente las mismas funciones y a partir de 2015 ha sido acompañado por el Teaching Excellence Framework. Ambos programas sirven para asignar fuertes sumas del presupuesto público en el sistema universitario británico (Ratcliffe, 2015).

⁸ La UAM fue pionera internacionalmente en la llamada “marea métrica”; sin embargo, lo hace utilizando su tabulador y no sigue en demasía los instrumentos actualmente en boga globalmente, como el número de citas, el factor de impacto, el h-index, etc.

⁹ Aparte del salario base, los profesores-investigadores de la UAM tienen la posibilidad acceder a cinco becas y estímulos: la Beca de apoyo a la permanencia, el Estímulo a la docencia e investigación, el Estímulo a la trayectoria académica sobresaliente, el Estímulo a los grados académicos y la Beca al reconocimiento de la carrera docente (UAM, 2013). En los tres primeros casos se utiliza el TIPPA para determinar si el candidato puede recibir esa beca o estímulo; en este sentido, el TIPPA no sólo sirve para asignar categorías laborales; lo utilizan igualmente para determinar compensaciones económicas.

Tabla 1. Primer tabulador (cualitativo) de la UAM

Categoría y nivel	Escolaridad	Docencia profesional	Investigación	Ejercicio	Otros
Asistente "A"	Título: licenciatura	Un año como auxiliar de profesor o investigador	Reporte de investigación avalado por el jefe de área cuando no haya tesis	Un año de ejercicio profesional como analista	
Asistente "B"	Título: licenciatura	Un año de experiencia docente	Tres reportes de investigación con arbitraje	Dos años como analista	
Asistente "C"	Título: licenciatura	Dos años de experiencia docente	Seis reportes de investigación con arbitraje	Tres años como analista	
Asociado "A"	Título: licenciatura	Tres años de experiencia docente Material didáctico para una U.E.A.	Un artículo con arbitraje	Cuatro años como analista	
Asociado "B"	Título: licenciatura	Cuatro años de experiencia docente Material didáctico para dos U.E.A. Un programa de U.E.A.	Dos artículos con arbitraje	Jefe de departamento	Cursos extracurriculares cuya suma sea de 20 horas
Asociado "C"	Título: licenciatura, 50% créditos de maestría	Cinco años de experiencia docente Material didáctico para tres U.E.A. Dos programas de U.E.A.	Tres artículos con arbitraje	Jefe de departamento. Encargado de proyecto (2 años)	Comités de carrera licenciatura Dirección de laboratorios Cursos extracurriculares por 30 horas

U.E.A.: Unidad de Enseñanza Aprendizaje.

Fuente: López Zárate *et al.*, 2000: 411.

legislación universitaria distingue entre 12 distintas categorías (algunas con hasta cuatro niveles), que suman un total de 33 "clases" —por llamarlas de algún modo— de académicos. Éstas van desde el escalón más bajo que

corresponde al "ayudante A" (que debe de tener un mínimo de 2 mil 200 puntos para aspirar al puesto), hasta el escalón más alto, que es el de "profesor de carrera C", que debe de tener un mínimo de 55 mil puntos.

Tabla 2. Fijación de categoría, de nivel, o de categoría y nivel del personal académico

Categoría	Nivel	Puntos
Ayudante	A	2,200
Ayudante	B	3,300
Ayudante de posgrado	A	6,600
Ayudante de posgrado	B	7,700
Ayudante de posgrado	C	9,900
Profesor asistente de carrera	A	6,600
Profesor asistente de carrera	B	7,500
Profesor asistente de carrera	C	8,600
Profesor asociado de carrera	A	13,200
Profesor asociado de carrera	B	15,700
Profesor asociado de carrera	C	19,100
Profesor asociado de carrera	D	24,000
Profesor titular de carrera	A	29,000
Profesor titular de carrera	B	40,000
Profesor titular de carrera	C	55,000
Técnico académico asociado de carrera	A	5,500
Técnico académico asociado de carrera	B	6,600
Técnico académico asociado de carrera	C	8,600
Técnico académico titular de carrera	A	13,200
Técnico académico titular de carrera	B	17,000
Técnico académico titular de carrera	C	22,000
Técnico académico titular de carrera	D	29,000
Técnico académico titular de carrera	E	40,000
Personal académico por obra determinada en áreas clínicas auxiliar	A	6,600
Personal académico por obra determinada en áreas clínicas auxiliar	B	8,600
Personal académico por obra determinada en áreas clínicas titular	A	13,200
Personal académico por obra determinada en áreas clínicas titular	B	17,000
Ayudante de tiempo parcial		3,300
Profesor asistente de tiempo parcial		6,600
Profesor asociado de tiempo parcial		8,600
Profesor titular de tiempo parcial		13,200
Técnico académico auxiliar de tiempo parcial		4,400
Técnico académico titular de tiempo parcial		8,600

Fuente: UAM, 2013: 151-152.

Existen dos versiones del TIPPA (muy similares entre sí): una que se usa para el ingreso de nuevos académicos a la universidad, y otro que se emplea para las promociones y becas de los académicos ya contratados. El tabulador de ingreso otorga puntos en más de 80 rubros

(donde cada rubro puede representar un producto terminado, un grado académico o una habilidad certificada) (Tabla 3).¹⁰ Por su parte, el tabulador para promoción otorga puntos en más de 110 rubros.

¹⁰ No deja de ser llamativo que la universidad utilice un procedimiento de evaluación esencialmente idéntico para contratar a un nuevo profesor (del que se puede saber nada o casi nada) que para evaluar a un profesor que lleva trabajando varios años o incluso décadas en la institución.

Tabla 3. Tabulador para Ingreso y Promoción del Personal Académico (TIPPA), UAM (fragmento)

		Mínimo	Máximo	Topes
1	Experiencia académica		50,000	
1.1	Docencia		30,000	
1.1.1	Impartición de cursos		9,000	
1.1.1.1	Cursos a nivel medio superior	100	100	900
1.1.1.2	Cursos a nivel licenciatura	330	330	7,200
1.1.1.3	Cursos a nivel posgrado	550	550	9,000
1.1.2	Elaboración de planes y programas de estudio		1,500	
1.1.2.1	Programas de licenciatura	220	220	1,500
1.1.2.2	Programas de posgrado	220	220	1,500
1.1.2.3	Plan de licenciatura	550	550	1,500
1.1.2.4	Plan de especialización	550	550	1,500
1.1.2.5	Plan de maestría	550	550	1,500
1.1.2.6	Plan de doctorado	550	550	1,500
1.1.3	Preparación de materiales didácticos		18,000	
1.1.3.1	Paquete didáctico (manual)	220	660	9,000
1.1.3.2	Notas de curso normal	220	660	9,000
1.1.3.3	Notas de curso especial	220	1,100	9,000
1.1.3.4	Antologías comentadas	110	660	5,400
1.1.3.5	Libros de texto	2,200	6,600	18,000
1.1.3.6	Documentales (audiovisuales, videos, cine, fotografías y diaporamas)	220	660	9,000
1.1.3.7	Equipo de laboratorio (modelos tridimensionales, diseño y construcción)	660	1,980	18,000
1.1.3.8	Desarrollo de paquetes computacionales	660	6,600	18,000
1.1.3.9	Traducciones publicadas de libros	110	660	5,400
1.1.3.10	Traducciones publicadas de artículos	20	110	900
1.1.3.11	Traducciones editadas de documentales	110	110	600
1.1.4	Dirección de tesis		6,000	
1.1.4.1	Licenciatura	220	220	6,000
1.1.4.2	Especialización	330	330	6,000
1.1.4.3	Maestría	440	440	6,000
1.1.4.4	Doctorado	880	880	6,000
1.2	Investigación		30,000	
1.2.1	Participación en programas de investigación comunicados idóneamente		27,000	
1.2.1.1	Reporte de investigación o técnico	110	330	5,400
1.2.1.2	Artículo especializado de investigación (artículo en revista, prólogo, introducción crítica, edición crítica de libro, o capítulo en un libro científico)	880	3,300	27,000

Fuente: UAM, 2013: 145.

Veamos ahora qué implica para los evaluados y para una institución como la UAM que se realicen las evaluaciones con un instrumento de esta naturaleza.

¿QUÉ LE POSIBILITA A LA INSTITUCIÓN UNIVERSITARIA HACER UNA EVALUACIÓN CUANTITATIVA?

Los métodos cuantitativos —como el TIPPA— sirven para legitimar los procesos institucionales de evaluación. Las medidas cuantitativas connotan objetividad y precisión; con información cuantitativa las decisiones dan la apariencia de ser menos “subjetivas” o “arbitrarias” (Carruthers, 2005). Estos métodos explotan la suposición generalizada de que las matemáticas son un “género de conocimiento preciso que no admite errores de apreciación o interpretación” (Navarrete *et al.*, 1976: 7). O, como lo expresaba una de las académicas de la UAM entrevistadas para esta investigación: el TIPPA “es un número exacto y eso da certidumbre” (E1).

Esto se ve con claridad en la justificación que se hace en la legislación universitaria de la UAM para explicar la introducción del tabulador cuantitativo de 1985:

La aplicación de este criterio [cuantitativo] permitió otorgar una respuesta idónea a los requerimientos de la comunidad universitaria en relación con el Tabulador, en virtud de que *una de las exigencias más solicitadas por dicha comunidad fue la de lograr un documento objetivo que eliminara las injusticias derivadas de los amplios márgenes de valoración que posibilitaba el Tabulador anterior* (UAM, 2013: 135. Énfasis añadido).

Por otra parte, en contraste con las evaluaciones cualitativas, las evaluaciones numéricas facilitan a las instituciones la comparabilidad entre sujetos evaluados (se comparan dos números y se observa cuál es el mayor y el menor). Por ejemplo, comparar y jerarquizar

en una lista de prelación a diez candidatos para una plaza utilizando criterios cualitativos, resulta más complicado que haciéndolo con criterios numéricos.

¿QUÉ POSIBILITA A LOS ACADÉMICOS SER EVALUADOS CON UN MÉTODO CUANTITATIVO?

Uno de los efectos que la evaluación cuantitativa tiene para los académicos es poder *commensurar* los distintos productos de su trabajo. La conmensuración implica usar números para crear relaciones entre objetos; convierte distinciones cualitativas en distinciones cuantitativas, por lo que las diferencias se expresarán como magnitudes determinadas siguiendo una medida establecida. Las diferencias expresadas de este modo se convierten en asuntos de *más* o de *menos*, y ya no en asuntos de calidad, especie, tipo o clase. Sea cual sea el área donde se utilice la conmensuración, ésta sirve como una estrategia tanto de integración como de distinción: establece una relación común entre objetos. Al hacer esto, se borra aquello que es distintivo de las cosas que son conmensuradas. La relación cuantitativa producida por la conmensuración hace posibles nuevas formas de distinción y borra otras (Espeland, 2002; Espeland y Stevens, 1998).

La conmensuración de los productos académicos es una práctica cognitiva cotidiana entre los profesores-investigadores de la UAM; en palabras de una académica: “un artículo vale cuatro cursos de licenciatura” (E3). Una afirmación de este tipo sólo puede cobrar pleno sentido visto a través de los ojos de la valoración numérica. Aquí se ve claramente cómo una distinción cualitativa (la diferencia entre ser docente de una clase de licenciatura y escribir un texto científico) se convierte en una distinción cuantitativa (la cantidad de puntos que se otorgan por dar una clase y los que se otorgan por publicar un artículo en una revista especializada). Si observamos los valores numéricos que el TIPPA establece para los

productos académicos, vemos que un curso de licenciatura vale 330 puntos y un artículo puede valer entre un mínimo de 880 y un máximo de 3 mil 300 puntos. Así, la idea de que “un artículo vale cuatro cursos de licenciatura” se hace transparente y comprensible. Dar cuatro clases de licenciatura vale mil 320 puntos (330 x 4), una cantidad que en efecto está dentro del rango entre los 880 y los 3 mil 300 puntos que son otorgados por un artículo científico.

Siguiendo esta lógica, se ha dado el caso¹¹ en que un académico de la UAM escribió un libro, pero antes de publicarlo como tal estimó más conveniente convertir el contenido del libro en cuatro artículos separados. Esta decisión —que puede resultar extravagante o hasta inverosímil visto desde otros principios de evaluación— tiene mucho sentido después de conmensurar el valor de un libro y el valor de cuatro artículos usando el TIPPA. Según éste, un “libro científico” vale entre 2 mil 200 y 6 mil 600 puntos; mientras que un “artículo especializado de investigación” vale entre 880 y 3 mil 300 puntos. Con esto la publicación de cuatro artículos oscila entre 3 mil 520 y 13 mil 200 puntos. De este modo, considerando que se obtuvieran los puntajes máximos de cada categoría, el mismo trabajo del libro, convertido en artículos, significaría obtener exactamente el doble de puntos (de 6 mil 600 a 13 mil 200 puntos).

En el contexto de otra “economía del prestigio”, tener un “libro de autor único” es considerado como uno de los logros más elevados (como sucede normalmente, por ejemplo, en el ámbito de las ciencias sociales y humanidades). Merecida o inmerecidamente, el valor simbólico de la autoría de un libro acostumbra ser altísima en los medios académicos e intelectuales. El nacimiento (publicación) de un libro acostumbra ir acompañado de una fiesta o ritual ceremonial (la “presentación del libro”). Los libros son valorados en más alta estima si vienen acompañados de una firma o

una dedicatoria del “autor”; son objetos que se obsequian, etcétera. Sin embargo, con la asignación de puntos del TIPPA, el libro se convierte en un bien cuyo valor se reduce a puntos; puntos que individualmente son idénticos a los que se obtienen por actividades tan disímiles como dirigir un programa radiofónico, estudiar un curso de idiomas o construir un equipo de laboratorio. Así, la conmensuración —con la subsecuente posibilidad que tienen los profesores-investigadores de comparar y calcular el valor numérico de sus productos— permite que la publicación de cuatro artículos individuales sea una *inversión académica* más redituable que publicar un libro.

Todo esto tiene múltiples consecuencias prácticas para la labor tanto de las universidades como de los profesores. En una encuesta que realizó la revista *Nature* a rectores y otros directivos sobre la necesidad de que las instituciones describan con claridad a los profesores cómo serán evaluados y qué elementos se van a considerar en ese proceso, recogieron la opinión del decano de una universidad que sostenía que no deseaba presentar una lista de las cualidades que buscaba en los candidatos a obtener una plaza con definitividad, porque eso “podría alentar entre los profesores el comportamiento de palomear la lista en vez de alentar la innovación” (*Nature*, 2010: 845). Aunque la revista critica esa actitud y demanda como esencial la transparencia en los criterios de evaluación del trabajo científico, la opinión de ese decano no es un sinsentido. El caso de la UAM y sus métodos de evaluación (que son instrumentos altamente transparentes) puede, en efecto, promover ese comportamiento de conformarse con palomear ítems en una lista de requisitos. Eso representa una ventaja para los evaluados, pues pueden adaptar sus tareas para manipular a su favor los puntajes de las evaluaciones.

Veamos un ejemplo. Al hablar de los formularios de las evaluaciones de clase que los

¹¹ Debo este dato al investigador Ángel Díaz-Barriga.

alumnos llenan en los cursos de cada trimestre —de cuyos resultados depende en gran medida la obtención de una “beca de docencia” que otorga la UAM a la mayoría de sus profesores—, uno de los entrevistados contaba lo siguiente:

Yo una vez hice un experimento: que mis evaluaciones [de los estudiantes] salgan perfectas. Y salieron perfectas. Todas las cosas, todos los puntos los fui llevando a cabo, tal cual, más allá de la calidad de mis clases. Por ejemplo, hay un punto que siempre me fallaba, le preguntaba a los alumnos: “¿Recibió asesoría?”. Les dije [a mis alumnos]: “les voy a dar asesoría a todos, a fuerza”. Los programaba una vez al mes y los llevaba de cinco en cinco. Les daba asesorías a todos, para que todos dijeran “sí tuve asesoría”. [...] Todos [los puntos de la evaluación] fueron así. Di el programa el primer día, nunca falté, usé métodos. Todos los puntos. O sea, era un experimento. Fue bueno el experimento, porque es la posibilidad de tener un esquema para sacar una evaluación. [...] Y la calificación subió. ¡El consejo divisional me dio una carta de distinción! (E12).

Conocer de antemano exacta y puntualmente qué parte del trabajo de un académico será evaluada, le permite a los evaluados actuar de modo estratégico —más allá de la calidad de sus labores— para asegurar una buena evaluación.

¿NÚMEROS U OPINIONES? EL TABULADOR Y SUS DIFERENCIAS CON OTROS MÉTODOS DE EVALUACIÓN

Utilizando el vocabulario de la teoría del constructivismo estructuralista sobre el capital cultural, se podrían contrastar los sistemas de evaluación basados en “juicios cualitativos” con el sistema basado en “cálculos cuantitativos” (como el de la UAM), diciendo que este

último privilegia al *capital cultural institucionalizado* (i.e., constancias y títulos académicos) sobre el *capital cultural incorporado* (i.e., la acumulación de cultura ligada al cuerpo por un proceso de interiorización) (Bourdieu, 2000). En otras palabras, la evaluación a través de tabuladores como el TIPPA le da una potestad absoluta a los documentos validados institucionalmente que certifican un logro o habilidad (i.e., títulos, certificados y constancias, todos necesariamente acompañados de sus respectivos sellos, apostillas, membretes y firmas). Así, sólo se reconoce como válido el capital cultural “académicamente sancionado y garantizado en forma legal mediante títulos”, que son certificados “de competencia cultural que confiere a su portador un valor convencional duradero y legalmente garantizado” (Bourdieu, 2000: 146). Esta primacía de los títulos y certificados va en detrimento de la posibilidad de emitir juicios sobre el individuo evaluado atendiendo a sus habilidades y capacidades culturales interiorizadas, sus “signos de competencia” y los “indicios apenas perceptibles” sobre sus capacidades (como la forma de hablar, los modales, la actitud, la disposición, la sensibilidad, los gustos), que en muchos contextos son cruciales para evaluar y seleccionar candidatos (Bourdieu, 1999: 133).

Esto marca un fuerte contraste entre el modo de contratar y evaluar académicos en una universidad por medio del tabulador, y el sistema prevalente en muchas instituciones del medio universitario angloparlante e instituciones de otras esferas culturales que duplican ese sistema:¹² en numerosas universidades estadounidenses, por ejemplo, el capital cultural incorporado es crucial para salir airoso en los procesos de contratación para una plaza de tiempo completo con posibilidad de definitividad (conocidas como *tenure track*). Los candidatos preseleccionados a una de esas plazas hacen visitas al campus (de uno o dos días completos) y realizan una o dos

12 Se hace esta comparación sin emitir ningún juicio sobre cuál de estos sistemas es “mejor” o “peor”.

presentaciones públicas (*job talks* y *teaching demonstrations*) y se entrevistan (lo mismo en entrevistas formales que en intercambios más “informales”, como tomar café y la hora del almuerzo) con los profesores que son potenciales futuros colegas, con personal de recursos humanos, con alumnos, con autoridades universitarias (como directores, decanos y rectores). Como recomienda un profesor a los aspirantes: “el candidato debe estar preparado, sobre todas las cosas, a amar todo lo que ve y a todas las personas que conoce; debe encontrar intereses comunes con tantos miembros del Departamento como sea posible, y ayudar a todas esas personas a observar las ventajas de contratar a *este* candidato en vez de a los otros” (Modern Language Association, 2015: s/p).

Estos procedimientos de evaluación —que son intensivos en *copresencia* y en interacciones cara a cara entre evaluado y evaluador, donde la identificación y valoración del capital cultural incorporado es permanente— difieren marcadamente del tipo de evaluación que se lleva a cabo en la UAM, que es *documentalmente intensivo* (lo que implica, entre otras cosas, un triunfo del sistema “credencialista” [Collins, 1989], donde se reduce el número de aspirantes a un trabajo no con base en su conocimiento o habilidades, sino en la obtención de títulos). La evaluación por medio del tabulador impone que los miembros de las comisiones dictaminadoras y evaluadoras pasen mucho tiempo frente a los expedientes y escaso o nulo tiempo frente a los candidatos. En el caso de becas, los evaluadores nunca interactúan con el evaluado durante el proceso de dictaminación; cuando se trata de promociones a veces se hacen entrevistas breves. En el caso de ingreso de nuevo personal académico —como lo manifiesta un dictaminador (E13)— la interacción cara a cara entre candidatos a una plaza y los dictaminadores puede ir de escasos 15 minutos a una hora (dependiendo, en parte, del número de aspirantes).

Otra característica llamativa del método de evaluación por medio del tabulador es la

enorme limitación que se impone sobre la posibilidad de que los dictaminadores utilicen su juicio informado. Tal como está diseñado el sistema, el criterio de los miembros de las comisiones evaluadoras se limita esencialmente a tres tareas:

1. Verificar si cada ítem que debe ser evaluado está colocado en la casilla correcta del tabulador, y si no lo está reasignarlo a otro casillero. Por ejemplo, se puede reclasificar lo que el evaluado consideraba un “artículo de investigación” (casillero 1.2.1.2 del TIPPA) como un “artículo de divulgación” (casillero 1.3.4).
2. Asignarle a cada ítem un puntaje dentro del rango establecido por el tabulador en el casillero correspondiente. Por ejemplo, un “artículo de investigación” no puede recibir un valor menor a 880 puntos o mayor a 3 mil 300; un “artículo de divulgación” no puede recibir un valor menor a 110 puntos o mayor a 330, etc.
3. Verificar que el documento que se presentó como prueba de cada actividad sea el adecuado (en esta tarea también intervienen abogados de la universidad, para comprobar la validez de un título académico, constancia, certificado, etcétera). Si la constancia que presenta un candidato para justificar un ítem en su expediente es considerada “no idónea”, entonces ese ítem es descartado y ese producto reeditaré en cero puntos. Esto quiere decir que, en esta lógica de evaluación, si un saber no está institucionalizado —en la forma de constancias, títulos, certificados— no existe ni puede ser evaluado. Así, vale más una constancia de “comprensión de lectura del portugués” que dominar el alemán o el chino y no tener ninguna certificación para avalarlo. O, como lo dice un entrevistado que participó en

una comisión dictaminadora de área: es indispensable que haya “un documento por cada actividad” (E13).

Si regresamos en este punto a la comparación con los métodos de contratación y promoción usualmente empleados en las universidades estadounidenses, podremos observar con claridad algunos contrastes entre los métodos cualitativos y cuantitativos. Por ejemplo, es llamativo que el sistema de la UAM no recurra nunca al uso de cartas de recomendación, que son una de las piezas clave para la contratación de profesores universitarios en Estados Unidos y que se solicitan en las convocatorias para candidatos que deseen competir por una plaza con opción a definitividad (un “paquete” de aplicación típico incluye un *curriculum vitae*, una carta de intención, textos de autoría propia —*writing samples*— y cartas de recomendación de tres expertos en el campo de especialización del candidato).¹³

Las cartas de recomendación —que muchas veces, y no sin razón, son vistas únicamente como manifestaciones del capital social del recomendado— son instrumentos cualitativos que permiten, a quienes las redactan y a quienes las leen, considerar un número indefinido de cualidades de la persona sobre la cual la carta hace referencia. Esto permite valorar, exaltar y detallar las cualidades de un candidato sin estar limitado a los elementos numerados en una lista y también escapar del peso que un tabulador le otorga, en términos numéricos, a un producto o habilidad determinado. Las cartas de recomendación —cuando hacen un balance entendido de lo que un candidato ha logrado o se espera que pueda lograr en un ámbito académico— son utilizadas para tener en el proceso de evaluación el *juicio informado* de algunos expertos.¹⁴

Para las evaluaciones de la UAM, en cambio, no se utilizan ni se solicitan cartas de recomendación (ni reseñas de los trabajos publicados por los evaluados), por lo que se prescinde de las opiniones y juicios informados de expertos en el campo de especialización del evaluado. Con esto, en la UAM se observa: 1) un monopolio de los documentos probatorios sobre cualquier otra forma de aprehensión de las habilidades del evaluado; y 2) una primacía de la magnitud numérica estandarizada y predeterminada sobre el juicio de los expertos en un campo de especialización.

Se debe enfatizar, sin embargo, que las cartas de recomendación y demás procedimientos cualitativos de evaluación no están exentos de problemas. Aunque no se las ha tomado como objeto de estudio de manera suficientemente amplia (considerando cuánto peso tienen en los procesos de evaluación en varios sistemas universitarios), numerosos estudios (Schmader *et al.*, 2007; Trix y Psenka, 2003; Madera *et al.*, 2009; Grote *et al.*, 2001) han mostrado cómo las cartas de recomendación pueden reproducir negativamente estereotipos de raza y género, sin contar que usualmente omiten características del recomendado que podrían ser relevantes para los evaluadores que reciben las cartas (particularmente, aquellas características que son poco halagadoras). Otros estudios señalan cómo en las entrevistas y presentaciones destinadas a la contratación, factores no profesionales como la apariencia física de los candidatos puede influenciar la opinión de los evaluadores (Lee, 2016; Lee *et al.*, 2016). Finalmente, investigaciones recientes sobre cómo funcionan los comités académicos de evaluación (Lamont, 2010; Posselt, 2016) señalan cómo los miembros de esos comités —aunque toman en serio su trabajo y actúan profesionalmente— llegan

¹³ Se pueden ver numerosos ejemplos en el “*job bank*” de la American Sociological Association (<http://asa.enoah.com/Job-Bank-Information>) o de cualquier otra asociación profesional.

¹⁴ Incluso para las evaluaciones de promoción de un académico en las universidades estadounidenses o canadienses (como las *tenure revision*) —donde los evaluadores de la institución han sido colegas del evaluado durante años— se suelen solicitar cartas (*review letters*, *external appraisals*), a entre cuatro y seis expertos, en las que deben detallar su opinión sobre los aspirantes a definitividad o a promoción.

a considerar de manera desigual a los candidatos según su nacionalidad, por ejemplo, o le dan peso a “cualidades evanescentes” (detalles sutiles, como el uso del lenguaje, que consideran que les permite identificar rasgos de originalidad, inteligencia o competencia) que no forman parte de los criterios formales y explícitos de evaluación.

Como veremos, el miedo a ser perjudicado por los desvíos y limitaciones de los métodos y procedimientos cualitativos es precisamente uno de los motivos por los cuales un instrumento cuantitativo como el TIPPA goza de amplia legitimidad entre los profesores de la UAM.

LA PEOR FORMA DE EVALUACIÓN, EXCEPTO POR TODAS LAS DEMÁS

Al considerar las opiniones de los profesores-investigadores de la UAM parece claro que la evaluación por medio de un tabulador cuantitativo parece haber tenido sobrado éxito en su objetivo de hacer más legítimas las evaluaciones ante los ojos de los evaluados. Aunque en las entrevistas y cuestionarios con los académicos se mostraron críticos ante varios aspectos de la actual forma de evaluación a través del tabulador, la inmensa mayoría opinó que no les gustaría ser evaluados bajo ningún otro método.

En palabras de un investigador del área de ciencias sociales que recuerda la transición del primer tabulador al actual:

Creo que [el tabulador] es el mejor instrumento. Si no estoy mal... la UAM fue la primera institución que hizo un tabulador numérico. Veníamos de un tabulador no numérico y [había] quejas [ante] la *incertidumbre* y la *arbitrariedad* de las [comisiones] dictaminadoras: “me ponen los puntos que se les da la gana”. ¡Ah!, entonces vamos a hacerlo cuantitativo para evitar la subjetividad... Yo creo que es un buen instrumento, acota la arbitrariedad (E8).

Esta misma opinión se puede ilustrar más detalladamente con lo dicho por un investigador del área de ciencias básicas, quien pese a tener serios cuestionamientos hacia el método de evaluación de la UAM, dice que no lo cambiaría por otra forma de evaluación. Primero sostenía que

...aquí para evaluarte un artículo lo más seguro es que ningún miembro de la dictaminadora entienda lo que yo estoy escribiendo... un trabajo que yo hice con un colaborador de [un país europeo], simplemente se publicó como un preimpreso; pero claro en [ese país] los preimpresos son bonitos, con una pasta bonita... Por ese artículo me dieron 3 mil 300 puntos... Por otra parte, en una publicación que hice también con otro colaborador, en una revista de primer nivel... [sólo] me dieron 1200 puntos. [Se trataba de una publicación] que yo puedo decir: es lo mejor que he trabajado, me costó realmente... [si me llegan a preguntar] “a ver, ¿qué hiciste en la vida?”: [hice esto], la revista es de primer nivel. Y cuando vi la evaluación que me pusieron, yo me sentí ofendido... He platicado con algunas personas. Los que evaluaron... a lo mejor no saben de las revistas; les digo: “¿no saben distinguir lo que es un preimpreso de una publicación?”. No noté dolo ni nada de eso, sino incompetencia... hay incompetencia en el interior de la gente que está ahí de dictaminadora (E10).

Después, este mismo investigador dijo: “respecto a otro país, es distinto... cuando yo les comento que en mi universidad se evalúa con puntitos: ‘¡es barbarie!’. Y estoy de acuerdo, pero tienen otros estándares y otras maneras de evaluar” (E10). Sin embargo, cuando el investigador reflexionó sobre la UAM en el contexto de las universidades mexicanas, dijo: “en mi experiencia, en lo que conozco en todos lados, lo que he vivido en todos lados en comparación con la UAM [éste es] el mejor método de evaluación que existe en México... ”

la UAM [emplea un] sistema de evaluación que tiene sus defectos, pero es el mejor sistema de evaluación” (E10).

Las repuestas de los investigadores en la encuesta resultan reveladoras sobre las fortalezas que ven en la evaluación a través del tabulador; en particular, ilustran de manera manifiesta su recelo por el uso político que se pueda dar si se emplean otros métodos de evaluación.

Una de las preguntas del cuestionario indagaba sobre qué tan dispuestos estarían a que se utilizaran métodos de evaluación donde los evaluadores estuvieran menos constreñidos por los puntajes del TIPPA y tuvieran mayor discrecionalidad para emplear su juicio informado. Las respuestas a esta pregunta resultaron categóricas. Como un académico afirmó elocuentemente: “ampliar los grados de libertad de las comisiones [tendría] el efecto de detonar las miserias humanas entre los académicos” (C6).

Otras respuestas a esa pregunta enfatizan cómo no se sienten inclinados a modificar ni flexibilizar el sistema de evaluación por una serie de temores ante lo que podría causar tener evaluadores que cuenten con mayor libertad de juicio. Entre estos miedos se encuentra el de la corrupción y abuso de los evaluadores. Así, fueron frecuentes expresiones de este tipo: mayor discrecionalidad de los evaluadores “daría pie a otro tipo de corrupción... es preferible que haya reglas y tabuladores claros y que no se puedan cambiar” (C1); “La evaluación con discrecionalidad da lugar a corrupción” (C12).

Los académicos consultados también mostraban una preocupación ante el uso político de las evaluaciones por parte de los grupos de poder dentro de la universidad. Uno de ellos, por ejemplo, rechazó la idea de darles más discrecionalidad a los evaluadores

...porque la evaluación dependería más de criterios personales de quienes estuvieran en determinado momento formando parte de

la comisión. Se politizaría más la participación en tales comisiones. Esta es un arma peligrosa que puede ser copada por grupos de poder contra otros grupos antagónicos (C5).

Otro sostuvo que

...las comisiones dictaminadoras se han convertido, en ocasiones, en mecanismo de premio o castigo para los profesores dependiendo de los amores o desamores de los distintos grupos políticos. Instrumentos que permitieran una mayor discrecionalidad a los miembros de las comisiones, podrían producir más pérdida de confianza (C14).

Uno más afirmó que el uso político de la evaluación se practica pese a los múltiples límites que el tabulador impone sobre los evaluadores:

Si con el sistema actual puede darse cierto grado de discrecionalidad en las comisiones dictaminadoras, eliminando los rangos y puntajes éstas se convertirían en los botines de los grupos de poder en la universidad. Incluso cabe señalar que curiosamente los miembros de las comisiones dictaminadoras obtienen mejores evaluaciones cuando son parte de las mismas que cuando dejan de serlo (C18).

En general, las opiniones recogidas entre los académicos muestran cómo éstos consideran que la evaluación cuantitativa sirve para cerrar las puertas a la “subjetividad”: “ Toda evaluación siempre es tamizada por la subjetividad de quienes la ejercen y los parámetros y los rangos acotan la intervención emotiva y anímica de los miembros de las comisiones dictaminadoras” (C7). Además de esa capacidad para restringir lo objetivo, varios profesores-investigadores subrayaron que la asignación de puntos por medio del TIPPA traía consigo valores positivos, como permitir evaluaciones más justas, transparentes y

exactas: “si con la interpretación subjetiva de las comisiones dictaminadoras la evaluación es muchas veces injusta, con la discrecionalidad sería peor” (C22); “aún con sus defectos, establecer rangos hace más justa la evaluación” (C9); “siempre es mejor ser evaluado con reglas claras y precisas y no depender del humor de quien evalúa” (C10).

Podríamos sintetizar la opinión de los académicos de la UAM sobre la forma en que son evaluados parafraseando la famosa frase de Winston Churchill sobre la democracia y decir que para ellos “la evaluación cuantitativa del tabulador es la peor forma de evaluación, excepto por todas las otras formas de evaluación que se han intentado”.¹⁵ La legitimidad que tiene la evaluación a través del tabulador entre los profesores de la UAM proviene de la certeza que dan los rangos cuantitativos del tabulador (con mínimos “inviolables”), de los límites que el tabulador impone en consecuencia a los criterios subjetivos de los evaluadores, y del diseño institucional que permite la selección de los miembros de las comisiones dictaminadoras, además de la existencia de instancias de revisión.

CONSIDERACIONES FINALES

El caso particular de los métodos de evaluación de la UAM sirve para hacer algunas consideraciones generales sobre los efectos, virtudes y limitaciones de la cuantificación del mérito académico.

En primer lugar, las “tecnologías del mérito” (como el TIPPA) con las que se mide el desempeño de los académicos reducen crecientemente el número de dimensiones consideradas relevantes para evaluar el desempeño académico a la vez que los instrumentos de medición se han hecho mayoritariamente unívocos. Recuperando una fórmula de Pierre Bourdieu —que utilizó para referirse a estudiantes en la competencia desenfadada

a través de exámenes nacionales estandarizados—, el tabulador coloca a todos los profesores-investigadores de la UAM bajo una “jerarquía única de formas de excelencia” (1997: 44). ¿Qué pensaríamos si en los Juegos Olímpicos todos los jueces que califican boxeo, gimnasia artística y clavados utilizaran las mismas tablas de evaluación para otorgar medallas en deportes tan distintos? ¿No sería poco razonable evaluar con un mismo criterio a un atleta que tiene por objetivo noquear a puñetazos a su contrincante y a otro que debe mover con gracia y elegancia un listón? Algo parecido es lo que sucede en la UAM con el TIPPA. Para poder atender sus licenciaturas y posgrados, la universidad necesita en su plantilla de profesores a abogados, administradores, agrónomos, antropólogos, arquitectos, biólogos, comunicólogos, diseñadores (gráficos e industriales), economistas, filósofos, físicos, geógrafos, hidrobiólogos, humanistas, informáticos, ingenieros (civiles, eléctricos, industriales, mecánicos y químicos), lingüistas, matemáticos, médicos, nutriólogos, politólogos, químicos, veterinarios, psicólogos y sociólogos. Todos ellos, con sus disímiles criterios y prioridades profesionales, tienen que someterse a la rígida férula de un tabulador único.

Aquí se aprecian con claridad las debilidades y fortalezas de una forma de evaluación centrada en un instrumento como el TIPPA: por un lado, éste reduce las múltiples actividades y productos en los que pueden materializarse los logros de un profesor-investigador a una sola forma de entender y valorar el trabajo académico. Abogados, biólogos, diseñadores y demás especialistas son todos sometidos a un estándar exclusivo (en el sentido de que tiene “fuerza para excluir”) de evaluación de la actividad profesional. En segundo lugar, un instrumento de evaluación rígido y normalizado —además de cuantitativo, reticular e impersonal— tiene la cualidad, difícilmente sustituible, de disminuir considerablemente el peso de los

¹⁵ La frase original de Churchill rezaba “La democracia es la peor forma de gobierno, excepto por todas las otras formas que se han intentado”.

critérios informales y “evanescentes” que muchas veces llegan a tener un lugar preponderante en las evaluaciones que son dejadas a criterio de los evaluadores. Esto es crucial en un medio informado por el recelo, la desconfianza, la suspicacia y el temor que tienen las personas ante el uso arbitrario, discrecional, interesado y parcial que hacen aquellos que están en posiciones de autoridad (como los miembros de las comisiones de evaluación, en este caso). Con todos sus defectos, las evaluaciones cuantitativas por medio de un tabulador, aumentan la certeza y la confianza hacia los resultados y procedimientos de una evaluación.

En tercer lugar, el discurso de la meritocracia se ha emparentado cada vez más con la llamada “objetividad mecánica”, que se basa en la multiplicación de estándares numéricos para establecer criterios de evaluación y de toma de decisiones que son usados por administradores para revestir su autoridad con reglas rigurosas e inflexibles (Porter, 1996). Las principales fuentes de legitimidad institucional de esas herramientas cuantitativas de evaluación son: primero, estar en línea con instrumentos y políticas de administración universitaria que han cobrado gran impulso a nivel internacional; y segundo, disminuir la discrecionalidad y “subjetividad” en los procesos de evaluación.

Finalmente, el tabulador no sólo cuenta (da puntos) las actividades académicas, también —y principalmente— define *qué*

es lo que cuenta (lo que puede ser considerado) como actividad académica evaluable. El tabulador es una entidad total; todo lo que está fuera de él es, por definición, irrelevante a los ojos de la institución. El tabulador anticipa al tipo de académico que va a evaluar, no es un “modelo de” sino un “modelo para” un tipo de académico. Las evaluaciones no sólo *miden*; al delimitar categorías y clasificaciones también *crean* un tipo particular de académico (Hacking, 1986). Como pasa con otros sistemas de medición, los tabuladores, índices e indicadores cuantitativos empleados para calcular los logros académicos no miden únicamente un fenómeno, también sirven para instituirlo —diciéndolo de otro modo, las mediciones del mérito académico son *performativas*—. Esto quiere decir que los sujetos que son medidos con estos instrumentos (los profesores e investigadores universitarios) seleccionan y adecuan sus actividades pensando en cómo serán sopesadas en el proceso de medición. La medición misma, en estos casos, sirve para crear el objeto medido (por ejemplo, los académicos seleccionan qué, dónde, cuánto y cuándo publicar teniendo en mente el peso numérico que esas publicaciones tendrán en sus próximas evaluaciones). Como han observado Espeland y Sauder (2007) invocando el concepto metodológico de “reactividad” (*reactivity*): las personas cambian su comportamiento como reacción a ser observadas, medidas y evaluadas.

REFERENCIAS

- ABBOTT, Alison, David Cyranoski, Nicola Jones, Brendan Maher, Quirin Schiermeier y Richard Van Noorden (2010), “Metrics: Do metrics matter?”, *Nature*, vol. 465, núm. 7300, pp. 860-862.
- ABOTES, Hugo (2012), *La medida de una nación. Los primeros años de la evaluación en México. Historia de poder y resistencia (1982-2010)*, Buenos Aires, Itaca/UAM/CLACSO.
- BOURDIEU, Pierre (1997), *Razones prácticas*, Barcelona, Anagrama.
- BOURDIEU, Pierre (1999), *Meditaciones pascalianas*, Barcelona, Anagrama.
- BOURDIEU, Pierre (2000), *Poder, derecho y clases sociales*, Bilbao, Desclée.
- BOURDIEU, Pierre (2008), *Homo academicus*, Buenos Aires, Siglo XXI.
- CARRUTHERS, Bruce (2005), “The Sociology of Money and Credit”, en Neil J. Smelser y Richard Swedberg (coords.), *The Handbook of Economic Sociology*, Princeton, Princeton University Press, pp. 355-378.

- CARSON, John (2004), "The Science of Merit and the Merit of Science: Mental order and social order in early twentieth-Century France and America", en Sheila Jasanoff (coord.), *States of Knowledge: The co-production of science and social order*, Londres, Routledge, pp. 181-205.
- COLLINS, Randall (1989), *La sociedad credencialista. Sociología histórica de la educación y la estratificación*, Madrid, Akal.
- CRONIN, Blaise y Cassidy Sugimoto (coords.) (2014), *Beyond Bibliometrics: Harnessing multidimensional indicators of scholarly impact*, Cambridge, MIT Press.
- DÍAZ-BARRIGA, Ángel (coord.) (2008), *Impacto de la evaluación en la educación superior mexicana. Un estudio en las universidades públicas estatales*, México, UNAM/ANUIES/Plaza y Valdés.
- ENDERS, Jürgen (2015), "Una 'carrera armamentista' en la academia: los rankings internacionales y la competencia global para crear universidades de clase mundial", *Revista de la Educación Superior*, vol. 44, núm. 176, pp. 83-109.
- ESPELAND, Wendy Nelson (2002), "Commensuration and Cognition", en Karen Cerulo (coord.), *Culture in Mind: Towards a sociology of culture and cognition*, Nueva York, Routledge, pp. 63-88.
- ESPELAND, Wendy y Mitchell Stevens (1998), "Commensuration as a Social Process", *Annual Review of Sociology*, vol. 24, pp. 313-343.
- ESPELAND, Wendy y Mitchell Stevens (2008), "A Sociology of Quantification", *European Journal of Sociology*, vol. 49, núm. 3, pp. 401-436.
- GALAZ Fontes, Jesús y Manuel Gil Antón (2013), "The Impact of Merit-Pay Systems on the Work and Attitudes of Mexican Academics", *Higher Education*, vol. 66, núm. 3, pp. 357-374.
- GARCÍA Salord, Susana (2013), "Del recuento curricular a la evaluación diagnóstica y formativa", *Perfiles Educativos*, vol. XXV, núm. especial, pp. 82-95.
- GROTE, Christopher, William Robiner y Allyson Haut (2001), "Disclosure of Negative Information in Letters of Recommendation: Writers' intentions and readers' experiences", *Professional Psychology: Research and Practice*, vol. 32, núm. 6, pp. 655-661.
- HACKING, Ian (1986), "Making Up People", en Thomas Heller, David Wellbery y Morton Sosna (coords.), *Reconstructing Individualism*, Stanford, Stanford University Press, pp. 222-236.
- HICKS, Diana y Paul Wouters (2015), "The Leiden Manifesto for Research Metrics", *Nature*, núm. 520, pp. 429-431.
- IBARRA Colado, Eduardo y Norma Rondero (2005), "Anexo 1. Periodización y caracterización de los modos de regulación del trabajo académico en la UAM (1973-2004)", en Manuel Gil Antón (coord.), *La carrera académica en la UAM. Un largo y sinuoso camino*, México, Universidad Autónoma Metropolitana, pp. 1-42.
- LAMONT, Michèle (2010), *How Professors Think: Inside the curious world of academic judgment*, Cambridge, Harvard University Press.
- LEE, Sun Young (2016, 12 de enero), "Are Good-Looking Men Predestined to Do Well in Academic Careers?", *Times Higher Education*, en: <https://www.timeshighereducation.com/blog/are-good-looking-men-predestined-to-well-academic-careers> (consulta: 15 de enero de 2016).
- LEE, Sun Young, Marko Pitesa, Madan Pillutla y Stefan Thau (2015), "When Beauty Helps and When It Hurts: An organizational context model of attractiveness discrimination in selection decisions", *Organizational Behavior and Human Decision Processes*, vol. 128, mayo, pp. 15-28.
- LLOYD, Marion Whitney, Imanol Ordorika y Roberto Rodríguez (2011), *Los rankings internacionales de universidades, su impacto, metodología y evolución*, México, UNAM, Cuadernos de Trabajo de la Dirección General de Evaluación Institucional 7.
- LÓPEZ Zárate, Romualdo, Óscar González Cuevas y Miguel Ángel Casillas (2000), *Una historia de la UAM. Sus primeros 25 años*, México, Universidad Autónoma Metropolitana.
- MADERA, Juan, Michelle Hebl y Randi Martin (2009), "Gender and Letters of Recommendation for Academia: Agentic and communal differences", *Journal of Applied Psychology*, vol. 94, núm. 6, pp. 1591-1599.
- Modern Language Association (2015), "Interviews, Campus Visits, Job Talks, and Teaching Demonstrations", en: http://www.mla.org/job_interviews (consulta: 9 de marzo de 2015).
- Nature* (2010), "Assessing Assessment", *Nature*, vol. 465, núm. 7300, p. 845.
- NAVARRETE, Manuel, Marcos Rosenbaum y Michael Ryan (1976), *Matemáticas y realidad*, México, Secretaría de Educación Pública.
- ORDORIKA, Imanol (2004), "El mercado en la academia", en Imanol Ordorika (coord.), *La academia en jaque. Perspectivas políticas sobre la evaluación de la educación superior en México*, México, UNAM/Miguel Ángel Porrúa, pp. 35-74.
- ORDORIKA, Imanol y Roberto Rodríguez (2010), "El ranking Times en el mercado del prestigio universitario", *Perfiles Educativos*, vol. XXXII, núm. 129, pp. 8-22.
- PORTER, Theodore M. (1996), *Trust in Numbers: The pursuit of objectivity in science and public life*, Princeton, Princeton University Press.
- POSSELT, Julie (2016), *Inside Graduate Admissions: Merit, diversity, and faculty gatekeeping*, Cambridge, Harvard University Press.

- RATCLIFFE, Rebecca (2015, 2 de noviembre), "The Teaching Excellence Framework: Can higher education up its game?", *The Guardian*, en: <http://www.theguardian.com/education/2015/nov/02/teaching-excellence-framework-university-tef-student-data-higher-education> (consulta: 18 de noviembre de 2015).
- RICKER, Martin (2015), "A Numerical Algorithm with Preference Statements to Evaluate the Performance of Scientists", *Scientometrics*, vol. 103, núm. 1, pp. 191-212.
- RONDERO, Norma (2007), "Impacto de las becas y los estímulos en la producción del trabajo académico en México", *Sociológica*, vol. 22, núm. 65, pp. 103-128.
- SCHMADER, Toni, Jessica Whitehead y Vicki H. Wysocki (2007), "A Linguistic Comparison of Letters of Recommendation for Male and Female Chemistry and Biochemistry Job Applicants", *Sex Roles*, vol. 57, núms. 7-8, pp. 509-514.
- SWARTZ, David (1997), *Culture & Power: The sociology of Pierre Bourdieu*, Chicago, University of Chicago Press.
- TRIX, Frances y Carolyn Psenka (2003), "Exploring the Color of Glass: Letters of recommendation for female and male medical faculty", *Discourse & Society*, vol. 14, núm. 2, pp. 191-220.
- Universidad Autónoma Metropolitana (UAM) (2003), "Criterios de dictaminación. Comisión dictaminadora en el área de ciencias sociales", *Semanario de la UAM*, núm. especial, 6 de enero, s/pp.
- Universidad Autónoma Metropolitana (UAM) (2005), "Criterios de dictaminación. Comisión dictaminadora en el área de ciencias básicas", *Semanario de la UAM*, núm. especial, 3 de enero, s/pp.
- Universidad Autónoma Metropolitana (UAM) (2013), *Legislación universitaria*, México, Universidad Autónoma Metropolitana.
- VAN NOORDEN, Richard (2010), "A Profusion of Measures", *Nature*, vol. 465, núm. 7300, pp. 864-866.
- VERA, Héctor (2005), "Cuándo, dónde y cuánto. El tiempo, el espacio y las medidas como problemas sociológicos", *Sociológica*, vol. 20, núm. 58, pp. 105-129.
- VERA, Héctor (2015), "La medición cuantitativa del mérito. El caso de la Universidad Autónoma Metropolitana", ponencia presentada en el XIII Congreso Nacional de Investigación Educativa, Chihuahua, Consejo Mexicano de Investigación Educativa, 16-20 de noviembre.
- VERA, Héctor (2016, 20 de enero), "El sistema educativo y la ubicuidad de la evaluación", *Nexos*, en: <http://educacion.nexos.com.mx/?p=132> (consulta: 20 de enero de 2016).
- WILSDON, James (2015, 9 de julio), "The Metric Tide: An agenda for responsible indicators in research", *The Guardian*, en: <https://www.theguardian.com/science/political-science/2015/jul/09/the-metric-tide-responsible-indicators-research> (consulta: 1 de agosto de 2015).