



RESU

resu.anuiés.mx

Revista de la Educación Superior 47 (187) (2018)



ARTÍCULO

Los cuatro jinetes de la evaluación: productivismo, reduccionismo, cuantofrenia y simulación

The four horsemen of evaluation: productivism, reductionism, quantophobia, simulation

Héctor Vera Martínez*

* Instituto de Investigaciones sobre la Universidad y la Educación, UNAM.
Correo electrónico: hectorvera@unam.mx

Recibido el 19 de marzo de 2017; aceptado el 24 de septiembre de 2018.

Resumen

Este artículo describe y analiza cuatro efectos adversos producidos por las políticas e instituciones de evaluación del trabajo de los académicos de las instituciones de educación superior en México. Estos efectos son: *Productivismo y conspiración contra la calidad*; los ritmos de evaluación y los tabuladores le dan más peso a la cantidad que a la calidad de los productos académicos, desalentando con ello los proyectos más ambiciosos y de largo plazo. *Reduccionismo y saturación de los evaluadores*; dada la desmedida carga de trabajo que involucraría evaluar detenidamente los expedientes a su cargo, los miembros de las comisiones evaluadoras optan por simplemente premiar a quienes demuestran tener un gran número de productos terminados. *Cuantofrenia y numerología*; en la evaluación se privilegian cada vez más los indicadores numéricos y hay una dependencia obsesiva con la medición. *Academia clientelista y simulación de la evaluación*; en muchos casos se mantienen evaluaciones opacas y tendenciosas que cumplen en apariencia con los procedimientos formales, pero sólo benefician a los miembros de ciertas camarillas.

Palabras clave: Evaluación, Productivismo, Reduccionismo, Cuantofrenia, Simulación.

Abstract

This article describes and analyzes four adverse effects produced by the policies and institutions designed to evaluate faculty members in Mexican universities. These effects are *Productivism and conspiracy against quality*; professors are evaluated very frequently and the institutional evaluating rubrics for ranking faculty give more weight to the quantity of the products evaluated than to their quality, this combined situation discourages the pursuit of large scale projects. *Reductionism and evaluators fatigue*; the massive amount of work needed to evaluate all the required files pushes evaluators to simply reward those who can deliver large numbers of completed products. *Quantophobia and numerology*; evaluation through quantitative indicators and obsessive reliance on measurement. *Crony academia and the simulation of evaluation*; persistence of opaque and biased evaluations, in which formalities are met but evaluators benefit mostly the members of their own cliques.

Keywords: Evaluation, Productivism, Reductionism, Quantophobia, Simulation.

Introducción

En este artículo se realizará un análisis crítico de las políticas y los métodos de la evaluación del trabajo académico en instituciones de educación superior (IES) en México. En primer lugar se hace una descripción panorámica de los problemas más visibles que presenta la evaluación académica. Posteriormente, se detallarán cuatro efectos adversos de la evaluación, ilustrándolos con las opiniones de académicos de varias instituciones públicas, obtenidas por medio de entrevistas semiestructuradas y encuestas con repuesta voluntaria. Estos efectos son lo que llamo los cuatro jinetes de la evaluación:

1. *El productivismo y la conspiración contra la calidad.* El sistema de evaluación premia una habilidad más que ninguna otra cosa: la capacidad de entregar productos terminados. Ser “productivo” se ha convertido en la cualidad más rentable en el medio académico. Esto, contrario a lo que se esperaba cuando se diseñaron las actuales políticas de evaluación, ha tenido el efecto no deseado de desalentar la investigación de calidad. Un efecto no deseado y no previsto de las políticas de evaluación es que vale más hacer “más cosas” que hacer “mejores cosas”.
- ii. *Reduccionismo y saturación.* Las actuales condiciones de evaluación hacen imposible una buena evaluación. Los procedimientos de evaluación (tanto en las IES como en los programas federales centralizados) sobrecargan de tal manera a los evaluadores que es materialmente imposible que utili-

cen su experiencia y sus conocimientos acumulados para aquilatar con el debido detalle los productos que deben evaluar. Ante tal situación, se ven constantemente arrinconados a premiar lo que el sistema pide: tener muchos productos terminados.

- iii. *Cuantofrenia y numerología*. La academia mexicana, como la de otros países, cada vez está más presionada para evaluar y rendir cuentas a través de indicadores numéricos. El fetiche de la cantidad ha llevado a la cuantofrenia (patología que consiste en una dependencia obsesiva de la medición) y a la numerología (una fe arcana en que los números esconden una esencia oculta de las cosas).
- iv. *Academia clientelista y simulación de la evaluación*. En varias IES hay una fuerte discordancia entre las reglas formales de evaluación (tal como están especificadas en la legislación y guías oficiales) y la forma efectivamente existente de evaluar (las prácticas y costumbres institucionales con las que en realidad se distribuyen las promociones, distinciones y sobresueldos). Esto hace que las evaluaciones sean opacas y tendenciosas. Además, fomenta la sospecha sobre las intenciones de las autoridades y los evaluadores, y que se vea como un mal menor a las formas de evaluación que promueven el productivismo y la cuantofrenia (pese a sus conocidos defectos).

Datos y metodología

La información empleada en este artículo fue recabada de cuatro fuentes principales.

En primer lugar, se analizaron y compararon los tabuladores de puntos y los reglamentos de ingreso, promoción y permanencia de treinta y nueve IES y centros públicos de investigación.

Segundo, se realizaron treinta y tres entrevistas semiestructuradas con profesores-investigadores de tiempo completo de cuatro universidades públicas mexicanas: Universidad Nacional Autónoma de México (UNAM), Universidad Autónoma Metropolitana (UAM), Universidad Autónoma de la Ciudad de México y Universidad Autónoma de Nuevo León (UANL). Las disciplinas de especialización de los entrevistados son: matemáticas, física, biología, fisiología celular, sociología, economía, administración y derecho.

Tercero, las entrevistas fueron complementadas con dos cuestionarios de respuesta voluntaria. Cada cuestionario constaba de quince preguntas y podía ser respondido por Internet. Un cuestionario fue enviado a profesores-investigadores que dan clases en cuatro licenciaturas de la UAM que pertenecen a la División de Ciencias Básicas e Ingeniería y a la División de Ciencias Sociales y Humanidades. El otro cuestionario fue enviado a investigadores

de un instituto de investigaciones perteneciente a la Coordinación de Humanidades de la UNAM. El porcentaje global de respuesta fue del 10.2 por ciento.

Por último, se recuperaron algunas declaraciones sobre los sistemas de evaluación realizadas por académicos mexicanos en eventos públicos.

El análisis de los reglamentos y tabuladores sirve para observar cuál es el marco procedimental e institucional que habilita y constriñe las acciones de evaluados y evaluadores. Por otra parte, se emplean los testimonios de los académicos — y se citarán frecuentemente sus palabras textuales — para entender cómo se sitúan los actores en su contexto de actividad. Los testimonios, pues, no son presentados aquí como “pruebas”; se trata, más bien, de expresiones extraídas del discurso de los mismos agentes donde articularon su visión sobre cómo funcionan los mecanismos que hacen operar al sistema de evaluación al que están sujetos y donde cristalizaron verbalmente cuáles son, desde su perspectiva, las consecuencias de dicho sistema.

La evaluación del trabajo académico

La evaluación ha impregnado cada rincón del medio educativo; esencialmente, todos los actores que tienen algo que ver con la educación son hoy sujetos de evaluación (Jornet, 2016; Vera, 2016). La forma, el propósito y los resultados de esa ubicuidad evaluativa han variado según los grupos, instituciones y niveles educativos. Para las IES y centros de investigación, el balance no ha sido mayoritariamente positivo.

Un raudal de investigaciones sobre los efectos de las políticas de evaluación en la educación superior ha dibujado un mapa de los vicios y consecuencias negativas que las evaluaciones traen consigo. Vale la pena mencionar, al menos brevemente, algunas de esas secuelas.

Uno de los problemas más acuciantes de la evaluación académica en las IES es que ésta ha quedado ligada a la distribución de compensaciones económicas que complementan los exiguos salarios de profesores e investigadores. Eso ha derivado en que no se evalúa para mejorar sino para gestionar sobresueldos. Esto sucede en muchos países, pero en México la cuestión es particularmente aguda. El porcentaje de recompensas “complementarias” para los académicos puede significar el doble de su salario base (mientras que en otros países los estímulos no rebasan más del diez por ciento).

Al mismo tiempo, la evaluación se ha burocratizado de manera creciente. Los evaluados viven la evaluación como un ejercicio repetitivo y tedioso de completar formularios y fotocopiar constancias. Esta burocratización ha aniquilado cualquier sentido de autoreflexión que pudiera (y debiera) tener

la evaluación. Evaluar no sirve como un camino para recapacitar sobre la propia práctica académica y sobre cómo perfeccionarla a la luz de los resultados de evaluación.

El sistema de evaluación, por otra parte, parece basado en la desconfianza recíproca entre las instituciones y los evaluados: las primeras usan la evaluación como instrumento de supervisión, y los segundos sospechan que la evaluación puede estar viciada por intereses extra académicos.

La evaluación en muchas IES se ha concentrado en la necesidad de responder a indicadores. Se demanda, sobre todo, que los profesores publiquen artículos de investigación. Aunque las universidades enfatizan que son tres sus labores sustantivas (docencia, investigación y difusión de la cultura), se premia de manera desproporcionada a la investigación. Como en otros países, “publicar o perecer” es la regla principal del juego. Y dadas las tendencias inflacionarias de los indicadores (lo publicado nunca parece suficiente), ahora es posible “publicar y perecer” (Broad, 1981: 1137). Esto ha promovido que los indicadores se conviertan en las metas a alcanzar, es decir, que se confundan los medios con los fines.

Los instrumentos de evaluación no se contextualizan para atender las necesidades geográficas, económicas y sociales de las distintas IES. Se mide a todos con el mismo racero, sin distinguir entre campo de conocimiento, condiciones institucionales, trayectoria del investigador, etc. Además, los mecanismos de evaluación no son examinados crítica y abiertamente por las autoridades y la comunidad académica con el fin de perfeccionarlos. En palabras de De Vries y Navarro: “[...] para la evaluación académica se pueden aplicar instrumentos no validados, guardar en secreto los resultados, no publicar los métodos de cálculo, no analizar los resultados, pero sí asignar recursos a partir del conteo o el puntaje” (2013: 4).

Estas consecuencias de los actuales métodos de evaluación no tienen un impacto homogéneo en todas las instituciones. Distintas políticas de evaluación y estímulos reducen o acentúan estos efectos, aunque pueden venir acompañadas de sus propias paradojas y contradicciones. Veamos con un poco más detalle, por ejemplo, el problema de que la mayoría de los programas de evaluación muestran predilección por la investigación, en desmedro de la docencia. Para paliar ese desbalance, en 1992 se creó el Programa de Estímulos al Desempeño del Personal Docente (ESDEPED), que beneficia a 34 universidades públicas estatales (UPES).

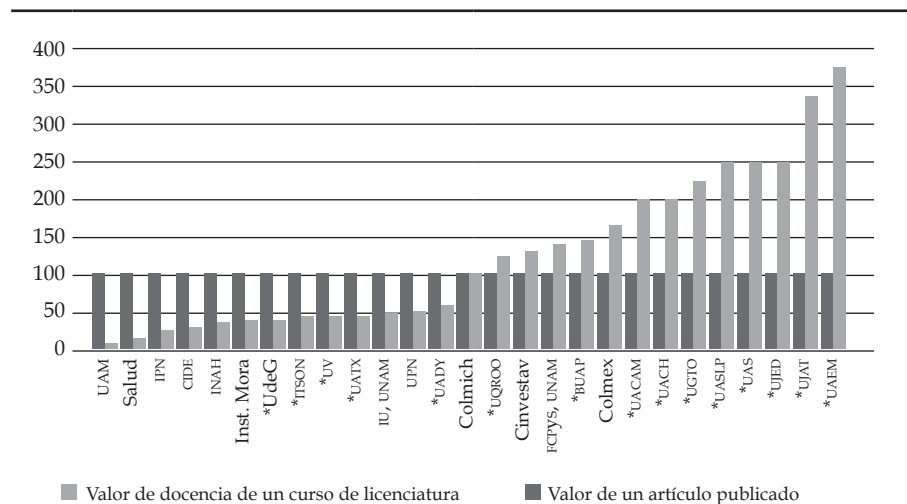
Comparando los instrumentos de evaluación de universidades que participan en el ESDEPED con los que se emplean en algunas universidades federales y centros de investigación, se puede ver que, en términos generales, el Programa ha funcionado para que la docencia de cursos de licenciatura esté

mejor valorada respecto a los puntos que se reciben por publicar artículos de investigación. Tras cotejar los tabuladores empleados en 27 IES y centros de investigación, para comparar el valor proporcional que le dan a la docencia de un curso de licenciatura con el puntaje que le asignan a la publicación de un artículo, se puede ver que, en promedio, en las instituciones que no participan en el ESDEPED un curso vale el 63 por ciento de un artículo publicado; en las UPES que participan en el ESDEPED, en cambio, un curso vale en promedio el 170 por ciento de un artículo publicado. Y de las 13 instituciones cuyos tabuladores otorgan más puntos a la docencia que a la publicación, diez pertenecen al ESDEPED.

La Gráfica 1 sintetiza esta información (en el Anexo 1 se detallan los datos de cada institución).

Gráfica 1
Valor proporcional de los publicados respecto a la docencia

Comparación del valor proporcional que tabuladores de 27 IES y centros de investigación le otorgan a la publicación de artículos y a la docencia de un curso de licenciatura



* UPES pertenecientes al ESDEPED.

Fuente: elaboración propia con base en los tabuladores y los reglamentos de ingreso y promoción de las IES enlistadas en el Anexo 1.

Ahora bien, todo esto no está libre de paradojas. Si observamos los criterios de asignación de recursos del ESDEPED (SEP, 2017), se puede ver que los montos se distribuyen entre las universidades de acuerdo a ciertos indicadores como: el número de profesores de tiempo completo que son miembros del Sistema Nacional de Investigadores (SNI); el número de profesores con reconocimiento de Perfil Deseable en el Programa para el Desarrollo Pro-

fesional Docente de la SEP; el número de publicaciones en revistas indizadas en la base de información bibliográfica *Scopus*; el número de programas en el Padrón del Programa Nacional de Posgrados de Calidad del Conacyt; etc. Todos esos indicadores están vinculados con programas e índices donde, nuevamente, se valora la publicación de artículos sobre todas las cosas. Así, hasta en un programa que busca apuntalar la docencia, viene escondida la tendencia de privilegiar la investigación. Incoherencias y contrasentidos como éstos abundan en la evaluación académica en México.

En lo que sigue me detendré en cuatro efectos de la evaluación (vinculados entre sí) que considero especialmente relevantes. Para detallar estos puntos – además del análisis de los tabuladores y reglamentos de evaluación empleados en las IES públicas mexicanas – se utilizarán las opiniones de académicos que he recabado por medios de entrevistas y encuestas.

I. Productivismo y conspiración contra la calidad

El científico británico Peter Higgs recibió el premio Nobel de física en 2013 por su explicación del origen de la masa de las partículas elementales (la así llamada “partícula de Dios”). En una entrevista Higgs hizo una declaración tan sorprendente como reveladora: “En el mundo de hoy no podría conseguir un trabajo académico. Tan sencillo como eso. No me considerarían suficientemente productivo” (Aitkenhead, 2013). Desde 1964, cuando dio a conocer su principal descubrimiento, Higgs ha publicado menos de diez artículos.

En México ocurren cosas similares. En octubre de 2016 el historiador Antonio García de León recibió – por su obra *Tierra adentro, mar en fuera* – el premio Clarence H. Haring, que otorga cada cinco años la American Historical Association al autor del libro más destacado de historia de América Latina. Ese mismo mes, García de León reflexionaba sobre las condiciones que le permitieron realizar esa investigación y lamentaba que “[...] todo conspira en contra de un trabajo largo. Los mismos sistemas influyen en eso. Yo debo confesar que para hacer este libro me tuve que salir del SNI; ahí no lo hubiera podido hacer [...]. A lo mejor para escribir el libro grande hay que salirse del sistema. Y soy realmente muy feliz de haberme salido” (García de León, 2016).

Este no es un caso aislado ni un comentario excéntrico. Las políticas de evaluación han traído consigo una inadvertida e indeseada *conspiración contra la calidad* que cala hondo en el medio universitario. En entrevista, un joven académico del Instituto de Investigaciones Jurídicas (IIJ) de la UNAM (donde todos los académicos tienen que presentar informes anuales a través un ta-

bulador cuantitativo llamado *Contabilidad de obligaciones anuales mínimas* [III, 2014]) expresó lo siguiente a propósito de las dificultades para imaginar y realizar trabajos ambiciosos que aspiren que a tener alguna trascendencia:

La libertad de cátedra me parece un valor fundamental, y creo que uno como investigador debe escoger, en ciertos momentos de su carrera profesional, la libertad de cátedra sobre la evaluación. O sea, uno debe correr el riesgo de decir “este no es un producto normal para evaluarse, pero vale la pena por estas y estas razones”. Entonces uno tiene que impulsarlo. [...] Por razones parecidas a las que [las evaluaciones por medio de] puntajes fraccionan las investigaciones de largo plazo, también podrían detener, en algunos casos, las investigaciones originales, porque los riesgos y las inversiones son mayores. [...] Cuando González Casanova escribió *La democracia en México*, no existían [las evaluaciones], entonces pudo hacer ese magnífico libro. Es más difícil para nosotros, porque los comités y los sistemas de reglas tienen la ventaja de que pueden hacer producir a los que están retrasados, pero frenan a quienes tienen proyectos de largo plazo.

Las investigaciones de largo aliento —aquellas que podrían tener un impacto más significativo en su campo respectivo— aparecen hoy como un riesgo laboral, como algo que se hace a pesar de las políticas de evaluación. Curiosamente, esas políticas fueron diseñadas con la intención de promover y premiar las contribuciones más importantes en cada campo de conocimiento; pero en la práctica producen (casi se podría decir que garantizan) *la mediana intelectual*.

Los sistemas de evaluación son operados —abierta o veladamente; reflexiva o inconscientemente— para que investigadores que usualmente producían poco se convirtieran en “obreros intelectuales”, suficientemente disciplinados para entregar de manera regular un número mínimo de publicaciones, asesorías de tesis, ponencias, etc. Pero para lograrlo se han creado engorrosas condiciones de control y vigilancia (manifiestas en los informes detallados y repetitivos que exigen distintas instituciones) que, por un lado, ahogan el tiempo de los investigadores, y, por el otro, condicionan una alta proporción del ingreso económico bajo la condición de que entreguen un número elevado de publicaciones. Esto ha fomentado un boicot en contra de los productos de alta calidad, que acostumbran ser lentos y que por naturaleza conllevan una alta posibilidad de fracaso (si hacer descubrimientos significativos fuera una apuesta sencilla y segura, todos seríamos grandes investigadores).

Sin embargo, no se debe pensar que los académicos están en contra de ser evaluados o que piensen que publicar artículos sea algo carente de mérito. Ven en la evaluación una utilidad potencial. Sobre lo que expresan inconformidad es la forma particular de evaluación a la que están sometidos —

aunque no sepan con claridad con qué la podrían sustituir —. Así se ve en el testimonio de un investigador del Instituto de Biología de la UNAM:

Todo tiene que ver con cuál es la base, tu línea basal [...]. Yo estaría de acuerdo con una evaluación de “estás mejorando” o “vas para abajo”, como se hace en cualquier negocio donde tienes que hacer dinero. Estas cosas son bien importantes. Por ejemplo, en el fútbol a los jugadores les miden todo: cuántos kilómetros están corriendo por partido, cuántos *sprints* se echó, cuántas pelotas ganó, cuántos pases buenos dio, si los dio para adelante o para atrás, cuántos “mano a mano” ganó. Entonces, en las cositas chiquitas perciben en qué están mejorando o quedándose en la base. Y mientras te quedes en la base o mejores, pues estás bien [...]. El problema es cómo cuantificas cuáles son los parámetros correctos. Yo creo que ahí falta mucha investigación: ¿cuáles son los parámetros correctos para evaluar a un investigador?

Los académicos entienden que evaluar es importante y que les puede ayudar a mejorar. Pero comprenden que las actuales evaluaciones no sirven para eso. Saben también que hay ámbitos de su actividad —la cual no ven reducida a la manufactura en serie de *papers*— que consideran esenciales, pero que están depreciados o son ignorados. El mismo investigador decía:

Hay una disciplina que se llama etnobotánica, que tradicionalmente se ha [visto] como “¡ah!, la etnobotánica es un cuento, nada más dices historias”; pero lo que han hecho los investigadores es que usan el conocimiento tradicional para levantar empresas, para conservar la biodiversidad local en algunas poblaciones; por ejemplo, los rarámuri en la Sierra Tarahumara. Hay investigadores que han rescatado la riqueza de quelites que tienen allá, les explican de qué se trata su riqueza, les enseñan a cultivarlos. Y ahora es un producto que llega a la ciudad de México como producto gourmet. ¡Ese tipo de cosas tienen un impacto social tan grande y tan necesario para México! A lo mejor no es un impacto científico mundial [...], pero que se conserve la diversidad vegetal de México, que se beneficie a una comunidad, que se establezca una cadena económica que no existía, [eso] no tiene precio. ¿Y quién lo va a evaluar? Ése es el problema que tengo.

Las políticas de evaluación no solamente han creado una lógica que no promueve la calidad académica, sino que tampoco reconoce las aportaciones extra científicas (cívicas, comunitarias, económicas, ambientales, etc.) de los proyectos de investigación.

Lo que sí se ha fomentado —inadvertidamente— es una explosión de prácticas de dudosa honorabilidad (por no hablar de la abierta deshonestidad, como el plagio, las falsas coautorías, etc.). Dado que los evaluados saben

que es muy probable que cuando sean evaluados el volumen de lo producido pesará más que la calidad (i.e., la originalidad, la innovación, el rigor, la pertinencia, etc.), entonces planifican su labor para maximizar la cantidad de publicaciones que pueden obtenerse en cada investigación. Es así como se ha multiplicado, por ejemplo, la “unidad mínima publicable” (*least publishable unit*), que consiste en fragmentar un trabajo de investigación en partes tan pequeñas como sea posible mientras sigan siendo aceptadas por los editores de las revistas científicas (Broad, 1981; Sindermann, 1982: 33-36; Pérez Tamaño, 1996).

II. Reduccionismo y saturación de los evaluadores

¿Por qué en el sistema de evaluación mexicano se valora más la cantidad que la calidad? No es porque los evaluadores creen que esa sea la forma más adecuada de sopesar el valor del trabajo académico o porque ignoren que tal problema existe. Sucede porque los evaluadores, dadas las condiciones en que operan, usualmente tienen poco margen para hacer otra cosa.

Cuando los académicos envían sus productos a evaluar lo hacen con la impresión de que los evaluadores no tendrán tiempo para leer los artículos o libros de los expedientes. Esta impresión es usualmente cierta, lo confirman quienes han pertenecido a comisiones dictaminadoras. Un sociólogo de la UAM retrata de este modo el escenario cotidiano al que se enfrentaba mientras fue miembro de la comisión dictaminadora de ciencias sociales de esa universidad:

Te horroriza ver los expedientes cuando estás ahí. [...] Tú estás muy abrumado porque todos los días, todas las sesiones que llegas a la comisión ves más. [...] Tú vas de una semana a otra, llegas un día y ves un librero lleno, llegas al otro y ves dos libreros llenos. ¡Entonces te entra una angustia! Además, la gente te empieza a presionar porque quieren cobrar, porque a partir de un cierto momento del año les dejan de llegar las becas a menos de que tú ya los hayas evaluado. Y además te llega un concurso [de oposición] e imagínate que te sea donde participan 20 personas, tienes que revisar 20 cajas.

La falta de tiempo para sopesar con cuidado los productos de los evaluados implica que los miembros de las comisiones dictaminadoras tengan que confiar en signos indirectos de calidad, como la abundancia y la novedad de las referencias bibliográficas, la extensión del texto, el prestigio de la casa editorial, la pertenencia a un padrón de publicaciones de calidad, la presen-

cia de referencias a trabajos escritos en idiomas diferentes al español, etc. En palabras del evaluador arriba citado:

Se ve mucho el tema de las notas al pie de página y una bibliografía que sea copiosa, actualizada. Si de repente no hay una sola nota al pie — así aunque no leyeras nada, nada más hojeando [el artículo] no vieras ninguna nota al pie y la bibliografía fueran cuatro textos —, ahí pues está medio raro. [...] *Se puede leer el abstract y conclusiones; eso es [a lo que llegan los dictaminadores] que hacen más.* Y bueno, el idioma; si alguna cosa está en inglés puede ser un indicador de que eso está interesante. [...] Pero en términos generales, la forma más superficial era: dónde se publicó, el idioma, la bibliografía, las notas al pie, [...] *el abstract y los resultados (resultados de algo, que se viera que había una cosa de investigación).*

Otro profesor de la UAM, un historiador que también formó parte de una comisión dictaminadora — aunque veinte años antes —, describía una situación muy similar, mientras se lamentaba de que la evaluación no se hace con el rigor suficiente: “Si me llegaba a mí como dictaminador un libro, hacía lo mismo que hacen todos: *en el mejor de los casos* empiezas a ver las notas al pie o ves la introducción y las conclusiones”.

Sumado a la escasez de tiempo y a la sobreabundancia de material, está el problema de la especialización. Entre los académicos entrevistados, la mayoría opina que los dictaminadores son, en general, personas competentes y profesionales; pero no están seguros de si habrá alguien que sea experto en su materia particular, capaz de juzgar adecuadamente su trabajo. Es una duda razonable.

Una comisión de seis científicos sociales (o seis humanistas o ingenieros, etc.) estará casi siempre escasa en *expertise* para evaluar a los aplicantes de toda una universidad o todo un sistema nacional (como el SNI). Se podrá argüir, por ejemplo, que un lingüista (digamos un especialista en lexicología maya) puede leer un artículo de otro lingüista (aunque trate sobre fonología del ruso) y formarse una opinión razonable sobre el texto. Pero eso supone, claro está, que hay tiempo para leer el artículo con cuidado.

Continuando con el caso de la UAM, si se hace un cálculo general sobre la cantidad de productos que tienen que ser evaluados, se puede constatar que opinar algo como evaluador — aunque sea mínimo — sobre cada *item* en los expedientes evaluados resultaría una empresa colosal. En el *Informe de actividades 2015* de esa universidad, se reportó que las comisiones dictaminadoras evaluaron 57 178 “actividades”, de las cuales 4 634 eran artículos especializados de investigación (UAM, 2016: 73). En la UAM hay nueve comisiones dictaminadoras, con seis integrantes cada una; tenemos entonces que en

promedio cada comisión tuvo que evaluar 6,353 actividades (incluyendo 514 artículos). Con eso, *cada dictaminador individual tendría que leer 85 artículos y evaluar otras 973 actividades* (que pueden incluir libros, patentes tecnológicas, ponencias escritas, etc.) en una faena que, se supone, los ocupa sólo un día por semana.

A estas condiciones de evaluación hay que sumar la exigencia de emplear indicadores numéricos para guiar el manejo de las universidades (Rotenburg *et al.*, 2015), que lleva a las autoridades a alentar que los investigadores entreguen cuentas constantes y sonantes. Por eso no es raro que el proceso evaluativo se adultere y quede reducido a contar cuántos artículos de investigación fueron publicados por cada investigador. Ese reduccionismo evaluativo cumple una función en el sistema; no es abiertamente reconocido por las autoridades, y tanto evaluadores como evaluados lo consideran poco recomendable, pero *sin el reduccionismo, el tiempo y la cantidad de trabajo para evaluar serían insostenibles*.

Haciendo referencia a esta situación, un investigador de un instituto de la Coordinación de Investigación Científica de la UNAM lamentaba que las comisiones dictaminadoras usen como único criterio los indicadores bibliométricos. Él ha dirigido convenios de colaboración con importantes entidades públicas que han sido alentados por Rectoría y por el director de su instituto, pero hacerlo le ha quitado tiempo para publicar. Y aunque la institución se jactó de esos convenios, cuando llegó la hora de evaluarlo individualmente fue castigado por tener pocos artículos publicados. En entrevista apuntaba con ironía que:

En el PRIDE [Programa de Primas al Desempeño del Personal Académico de Tiempo Completo] me bajaron porque dicen que no tengo el número de publicaciones, que *deben ser dos por año*. Yo les digo: “miren, esta publicación es posible partirla en dos. Lo hice uno grande en lugar de dos chiquitos.” Uno no puede discutir. Dicen que sí, pero al momento ni lo ven, no leen nada, no son de la materia. [...] Cuando hablo con colegas, les digo: “[...] *mi hija de la primaria puede hacer esta evaluación*. No hay ningún problema, *nada más tiene que saber contar hasta diez*.” No hacen una evaluación. [...] Uno se tiene que preguntar si las publicaciones y citas son los resultados verdaderos de la ciencia o si constituyen más bien el discurso de la ciencia. Existe el riesgo de que nuestra discusión se concentre en lo que se puede medir fácilmente –en el número de publicaciones y citas– en lugar de lo que realmente importa.

Sólo hace falta “saber contar hasta diez”. Ésta es una hipérbole, pero lleva consigo una letal dosis de veracidad.

III. Cuantofrenia y numerología evaluativa

No sostendré aquí que contar, calcular y cuantificar sean, por sí mismas, prácticas nocivas para realizar una valoración. Son dañinas sólo si están mal hechas, lo que sucede frecuentemente. La cuantificación es un instrumento intelectual poderoso y necesario. Es una de las formas más efectivas para recabar y organizar grandes cantidades de información, para comparar entidades que parecerían incompatibles. Pero la cuantificación no es una herramienta infalible, ni su uso es siempre pertinente, pues puede ser también una estrategia cognitiva fatalmente reduccionista.

Costumbres poco inteligentes de evaluar las creaciones intelectuales han existido siempre. En 1862, en una carta a su amigo Engels, donde hablaba de sus avances para una segunda parte de la *Crítica de la economía política*, Karl Marx se quejaba diciendo: “[...] estoy expandiendo este volumen, pues esos zafios alemanes estiman el valor de un libro en términos de su capacidad cúbica” (citado en Wheen, 2006: 32).

Ese fetiche por la cantidad es patente también en la academia mexicana y en sus formas de evaluación, y tiene consecuencias palpables que tergiversan el sentido mismo de las labores sustanciales de los investigadores. Por ejemplo, como sostiene una científica del Instituto de Fisiología Celular de la UNAM:

Creo que la tendencia, no solamente en la UNAM sino en todo el sistema de evaluación científico —que incluye el SNI—, es evaluar a la gente basado en el número de publicaciones que tiene, lo cual creo que es un error bastante grave. [...] Ve, un premio Nobel tiene ochenta publicaciones, ¡ochenta! [El conocido investigador mexicano Z] tiene ochocientas publicaciones, [...] ¡pero es porque tiene citas en contra! ¡Así no es cómo se hace ciencia!

Hace falta, pues, preguntarse: ¿qué se gana y qué se pierde cuando se utiliza la cuantificación como el camino para conocer y evaluar algo? En términos generales, con la cuantificación se gana claridad y la simplicidad en la información, facilidad para comunicar resultados y eficiencia para comparar y categorizar objetos. Entre las pérdidas se puede mencionar que las estimaciones cuantitativas descontextualizan la información, acostumbran ser unidimensionales, pueden magnificar diferencias insignificantes (a veces con graves consecuencias para los evaluados), tienden a naturalizar criterios arbitrarios de evaluación (le dan la apariencia de hechos sólidos a lo que sólo son apreciaciones particulares de individuos específicos), pueden producir una engañosa impresión de precisión y ocultan cualidades valiosas pero difíciles de medir.

En México, entre el productivismo (que premia tener grandes cantidades de productos terminados) y el agobio de los evaluadores (que por la falta de condiciones adecuadas terminan sólo sumando el número de artículos publicados), la evaluación cuantitativa ha sido vista como una solución para varios problemas. En particular, se ha vuelto popular el uso de tabuladores que asignan puntos a los ítems que se deben evaluar.

El avance de los tabuladores como herramienta de evaluación ha sido consistente. Es difícil determinar con exactitud cuántas IES los utilizan —al menos formalmente—, pero cuando menos medio centenar de ellas lo hacen. Y no se trata de instituciones periféricas. Varias de las universidades más importantes del país los usan: tanto federales (como la UAM, el IPN y la UPN) como estatales (al menos 30 UPES dicen evaluar con tabuladores). Haciendo una estimación conservadora, calculo que las IES que utilizan algún tipo de tabulador cuantitativo en sus evaluaciones concentran cuando menos al 45 por ciento de la matrícula nacional de estudiantes de licenciatura en instituciones públicas (sin contar a la UNAM, donde los tabuladores se utilizan en algunas entidades y varias más están considerando adoptarlo).

Dada la lógica que impone en las evaluaciones, el tabulador es un instrumento al que no se le ha prestado suficiente atención. García Salord (2013 y 2010) ha mostrado que el *currículum vitae* es una herramienta clave para materializar el “recuento de curricular” (nuestra forma imperante de evaluación), pues éste se centra en enlistar un reducido número de productos del trabajo y le resta valor a las prácticas y los productos excluidos del formato del CV. Pero el tabulador cierra la pinza de ese recuento curricular, pues con él se asignan cantidades prefijadas de puntos a los productos enlistados en el CV. Con esos puntos se le da un peso específico a cada *ítem* considerado en la evaluación global; y con la sumatoria de puntos de cada expediente es sencillo determinar con “precisión” si se llegó a la meta mínima para alcanzar un estímulo, o comparar los “méritos” de todos los evaluados.

Por otra parte, si se observan con detalle los tabuladores de varias IES, es posible encontrar muchos ejemplos de cómo el sueño de la cuantificación produce monstruos. En los años 50, Pitirim Sorokin advertía que las ciencias sociales habían sucumbido ante la *cuantofrenia*: la patología de depender obsesivamente de la medición y de aplicar procedimientos matemáticos a fenómenos que no son aptos de ser matematizables (Sorokin, 1956). Desde hace varios años la evaluación académica en México padece de esta patología. Si se analizan los instrumentos de evaluación empleados en muchas IES, es posible encontrar una amplia gama de procedimientos cuantitativos diseñados para ponderar los méritos académicos, pero que, en realidad, tergiversan cualquier sentido de calidad.

Uno de los problemas más frecuentes de la cuantificación malhecha es concentrarse en medir lo que resulta fácilmente medible, en vez de medir lo que es necesario medir (que muchas veces es elusivo). ¿Qué elementos deberían considerarse, por ejemplo, para evaluar la calidad de un libro académico? Ciertamente sería complicado medir con precisión sus cualidades centrales: originalidad, contribución para aclarar alguna vieja incógnita, facilidad para comunicar ideas complejas, etcétera. Pero sería sencillo –y de una vacua exactitud– contabilizar algunas de sus características más superficiales, como el número de páginas o la cantidad de ejemplares impresos. ¿Qué se debería hacer en una evaluación: calificar de manera aproximada lo sustancial o medir con precisión lo irrelevante? En la numerología evaluativa mexicana se opta por lo segundo.

En los más recientes *Criterios específicos de evaluación* del área v del SNI, se señala –en su sección de “Criterios, características y calidad de los productos”– que las reimpressiones de libros se consideran como “indicador de impacto” de una obra y se evalúa el tiraje de las ediciones agotadas. Un dictaminador del área IV del SNI comentaba que observó entre sus colegas de la comisión evaluadora la costumbre de considerar el tiraje de un libro como un signo de su relevancia académica. Los evaluadores –decía– opinaban que un libro serio debía tener al menos mil ejemplares impresos; los libros electrónicos o los tirajes bajo demanda (que se pueden hacer en tandas de pequeños números, según se vayan necesitando) despertaban sospecha.

Para muchos, los “mil ejemplares” es una suerte de unidad de medida del mérito científico o un número mágico que transforma las ideas contenidas en las hojas impresas en algo que merece ser premiado. Por ejemplo, en el tabulador que se emplea en la Facultad de Ciencias Políticas y Sociales de la UNAM para la contratación de nuevos profesores, se indica que si un libro llega a una segunda edición, por cada mil ejemplares aumentará el 5 por ciento “de su valor en investigación” (FCPYS, 2008: 14).

No hace falta decir que el número de ejemplares impresos casi nunca refleja algo intelectualmente importante de un escrito. Desde hace tiempo los estudios sobre historia y sociología de la lectura han mostrado que la abundancia de libros no significa que ese material impreso tendrá algún impacto ni dice nada sobre cómo será utilizado por los lectores (Chartier, 1995). Son conocidos, además, casos de obras que durante décadas no vendieron más de un centenar de copias y luego fueron reconocidas como obras maestras. Y dado que la ciencia se mueve a través de sub-áreas muy especializadas, hay muchos productos de investigación que pueden tener un impacto considerable en una comunidad específica, aunque sólo sea de interés para 200 o 300

colegas (una impresión de mil ejemplares para esos trabajos significaría un uso ineficiente de los recursos editoriales de una institución).

Pero entre los vicios de la cuantofrenia, el tabulador de la Universidad Autónoma Chapingo es de los que lleva las cosas más lejos. Ahí, los libros de autor —“con arbitraje e ISBN” — sólo se consideran si tienen “un mínimo de 50 páginas y un tiraje mínimo de 100 ejemplares” y se contabilizan en la evaluación otorgado “0.5 puntos por página” (UACH, 2013: 21). Si los ingenieros agrónomos que diseñaron ese instrumento de evaluación fueran críticos literarios, dirían que los cuatro tomos de *Cincuenta sombras de Grey*, de E. L. James, son quince veces más valiosos que *La metamorfosis*, de Kafka.

IV. Crony academia y la simulación de la evaluación

Con todo lo reprochable que pueda parecer este panorama de las evaluaciones, hay una situación aún peor: cuando las evaluaciones no son defectuosas sino *simuladas*. Lo que más daña a largo plazo el funcionamiento del sistema es que hay instituciones que reproducen el tipo de evaluación y de distribución de prebendas característico en las universidades hasta la década de 1980. Hasta ese momento, lo cotidiano era que los directivos decidieran todo discrecionalmente: a quién contratar, a quién promover, a quién darle becas, etc. Reinaban casi con descaro el nepotismo, el compadrazgo y la sumisión ante la autoridad. Se trataba de una academia “clientelista”, una *crony academia* (semejante al *crony capitalism*, donde el éxito en los negocios depende del intercambio de favores y lealtades entre empresarios y funcionarios). Ese patrimonialismo académico no ha desaparecido, pero se le ha contenido y parcialmente remplazado con rígidos mecanismos burocráticos.

Hoy, sin embargo, atestiguamos algo que parece más tóxico: una combinación de lo peor de ambos mundos. En algunas instituciones los directores otorgan dádivas y distribuyen los recursos discrecionalmente, pero disimulan esa potestad bajo la apariencia de un sistema de evaluación imparcial. Se puede ilustrar esto con lo que sucede, por ejemplo, en la UANL (la segunda universidad que recibió más recursos del fondo del Programa de Carrera Docente en UPES, en el ejercicio fiscal 2016).

A través de su “Programa de estímulos para personal académico”, la UANL otorga a algunos de sus profesores de tiempo completo un estímulo con montos monetarios variables según los resultados en una evaluación cuantitativa. Los evaluados que acuden a la convocatoria del programa llenan la información sobre sus actividades (docencia, participación en congresos, etc.) en una plataforma electrónica y juntan una carpeta con evidencias. Formalmente, la

categoría a la que entran está determinada por el número de puntos que reciben en la evaluación. Una puntuación entre 301-400 corresponde al nivel I, lo que significa recibir un sobresueldo de un salario mínimo mensual; una puntuación entre 951-1000 corresponde al nivel IX (el más alto), que se traduce en 14 salarios mínimos (UANL, 2016: 5).

El principal problema con este procedimiento no es que sea cuantitativo o que premie el productivismo, sino que los evaluados no saben cómo se otorgan los puntos que determinarán el resultado. No hay en la UANL un tabulador público, que puedan consultar los evaluados, donde se indique cuántos puntos se le otorgan a cada actividad. Tampoco se sabe quiénes son las personas que valoran cada expediente.

Al preguntarle sobre esta situación a un profesor de la Facultad de Filosofía y Letras de la UANL, comentó:

Asumimos, bueno, yo así lo asumo, que no hay correlación [entre tu evaluación y] tu producción. [...] Colegas que no tuvieron una producción ni siquiera muy cercana a la mía están por encima, o muy por encima [en los estímulos]. [...] Si fuera por mi producción, yo estaría mucho más alto. Por ejemplo, aquí en la Facultad somos muy pocos SNI, pero muy pocos, y yo soy uno de ellos. [...] Una producción amplia no se refleja en los estímulos.

En su opinión, las evaluaciones sólo se usan como “control político”, pues presume que es el director quien asigna los recursos: “No hay transparencia; entonces [dependiendo] cómo te portaste en el semestre, te otorgan los niveles”. Narró incluso un caso que, a sus ojos, confirma el total divorcio entre mérito y recompensa que priva en su institución: en una ocasión una profesora “no participó en la convocatoria [del programa de estímulos] y aun así le asignaron un nivel, ¡incluso superior al que tenía antes!”

Este contexto de simulación y discrecionalidad fomenta tal desconfianza y recelo entre los profesores, que estar en los niveles más altos de los estímulos no es algo que dé buena reputación. En las palabras de este profesor: “No [da prestigio]. Nunca sabes ni dónde están. Al contrario, mucha raza lo esconde [...] Uno empieza a pensar cosas malas. Siempre te están diciendo: ‘este vato no hace nada’ — porque en realidad no hace nada — ‘y está ganando 20 mil pesos de estímulos’ ”.

Es difícil saber qué tan generalizados son estos vicios en otras universidades públicas del país. Pero un dato que podría ser indicativo es que de las de 34 UPES que reciben fondos del Programa de Carrera Docente (U040), sólo 18 por ciento de ellas hacen público el tabulador con el que asignan los puntos para la distribución de los estímulos.

Situaciones como éstas, sean vividas en carne propia o conocidas de forma indirecta, crean una *aversión por la opacidad*. En el fondo, el temor a la arbitrariedad justifica la existencia —y alienta la expansión— de los sistemas de evaluación inflexibles, estandarizados y cuantitativos, pues éstos son vistos como una forma preferible de hacer las cosas. El productivismo y la cuantofrenia, comparados con la opacidad y la simulación, se conciben como un mal menor.

La academia clientelista y las evaluaciones simuladas tienen así un efecto profundo en todas las instituciones, incluso en aquéllas donde las evaluaciones son reales y se apegan al reglamento. Ese efecto consiste en aceptar los defectuosos modos de evaluación burocráticos e inflexibles, pues se les considera como el “mal menor”.

Esto se puede ilustrar vivamente con lo dicho en entrevista por una investigadora del III, UNAM (donde, como ya se mencionó, se evalúa a través de puntajes mínimos anuales). Al preguntarle si preferiría —en lugar del tabulador cuantitativo— ser evaluada por un comité de expertos en su campo que deliberara libremente y tuviera manga ancha para emitir sus juicios, respondió:

¡Se me haría nefasto! [...] Son sistemas que [...] abren la arbitrariedad y la subjetividad. [...] Además, hay muchísima incertidumbre. Ahí yo no sé cuántos puntos obtuve: de pronto me pusieron aquí, de pronto me pusieron acá, ¿y con base en qué? Y, además, si es así como me das el resultado, ¿cómo te digo que realmente no eres especialista, que evaluaste y no entendiste? Se me hace obscuro, se me hace arbitrario, se me hace subjetivo. A mí me encanta esto [de los puntajes mínimos], de que yo sé exactamente cuánto vale lo que hice y puedo darme una idea conforme va pasando el tiempo de que ya tengo los puntos que estoy [buscando], y que tengo expectativas buenas, mínimas, a base del “estado de derecho”. Yo sé desde antes cómo se me va a evaluar, cuáles son las normas con las que se me va a evaluar. Puedo planear mi año para generar los puntos que necesito; y cuando llega el número, sé de dónde viene. [...] Tengo una expectativa que se funda en razones.

Comparando su experiencia de las evaluaciones cuantitativas de su institución con lo que sucede en el SNI, considera preferibles las primeras:

¿Qué es lo que pasa en el SNI cuando conoces a la gente que está y que son expertos? De pronto dices: “¡No puede ser!, fulanito de tal tiene Nivel Candidato, zutanito tiene Nivel I y zutanito Nivel II, ¡¿pero cómo?!” Y cuando platicas con la gente que está en la comisión, tienen que evaluar doscientos [casos] cada día, una cantidad absurda; [como] si tuvieran el

tiempo infinito para – con su sabiduría sensacional y su imparcialidad total – leer y leer con cuidado.

Esta suspicacia hacia el SNI no es gratuita, pues ese programa no ha logrado despojarse de la sombra de la academia clientelista. En un estudio reciente (Frixione *et al.*, 2016) con una muestra aleatoria de investigadores del área de Humanidades y Ciencias de la Conducta del SNI, se comparó el nivel al que pertenecen los investigadores (I, II y III) con su producción científica demostrable. E hicieron lo mismo con todos los dictaminadores del SNI de esa misma área. Los resultados mostraron que hay una marcada disparidad en la productividad de los investigadores dentro cada nivel. Y, además, un cálculo indicó que 36 por ciento de los miembros de la comisión dictaminadora y 53 por ciento de los investigadores de la muestra aleatoria *no satisfacen los propios criterios establecidos por el SNI para estar en el nivel en el que se encuentran*. Encontraron, incluso, que hay investigadores en los tres niveles (aunque principalmente en el I) que tienen tan pocos productos de cualquier tipo, que es difícil explicarse cómo fue posible que resultaran seleccionados como miembros del SNI (Frixione *et al.*, 2016: 19).

Conclusiones

Los cuatro efectos adversos de la evaluación de académicos universitarios que se han analizado aquí están frecuentemente interconectados y se refuerzan mutuamente. El reduccionismo, por ejemplo, emplea frecuentemente a la cuantificación como uno de sus métodos predilectos para ahorrarles tiempo e inconvenientes a los evaluadores. Juntos, estos efectos forman parte de un extenso sistema de actores, instituciones y convenciones. No se trata de un sistema coherente, ni fue planeado para funcionar como una totalidad homogénea, pero sí es un sistema cuyas partes están vinculadas y donde el cambio en una de las partes conlleva cambios en las otras.

Que los procedimientos de evaluación en las IES mexicanas han traído consigo numerosas consecuencias adversas no es un secreto para los académicos ni para las autoridades universitarias. Entonces, ¿por qué persisten? En parte, porque la presente forma de evaluación ha cumplido, pese a todo, con algunos de sus objetivos manifiestos. Por ejemplo, en 1985, cuando la UAM reformó su Tabulador para Ingreso y Promoción del Personal Académico para hacerlo plenamente cuantitativo, se arguyó que con el cambio se esperaba que este nuevo instrumento eliminaría “las injusticias derivadas de los amplios márgenes de valoración” que el previo tabulador cualitativo

permitía (UAM, 2013: 135). Esto se ha logrado con creces. Cuando funcionan según fueron diseñados (es decir, cuando no sirven sólo para simular una distribución clientelar de los recursos), los nuevos métodos de evaluación le restan poder a los evaluadores y permiten mayor predictibilidad sobre el resultado de las evaluaciones (algo que los evaluados consideran sumamente valioso y que ha servido para legitimar a las instancias evaluadoras) (Vera, 2017a).

A pesar de eso, el productivismo, el reduccionismo y la cuantofrenia persisten no porque se considere que sean aspectos pertinentes de la evaluación —se sabe que no lo son—. Perduran y se diseminan porque cumplen importantes funciones “latentes” (es decir, son producto de consecuencias no intencionales y no advertidas de las acciones de los actores; consecuencias que le permiten al sistema institucional resolver problemas y continuar su operación habitual). La cuantificación y el reduccionismo de las evaluaciones no se diseñaron manifiestamente para ahorrarles tiempo a los evaluadores abrumados; pero ha cumplido de manera efectiva esa crucial función.

Por último, vale la pena mencionar aquí un tema poco atendido por los investigadores. A pesar de que se encuentran en medio de una maraña de prácticas e instituciones evaluativas que los limitan y constriñen, los académicos no son víctimas pasivas condenadas a vivir bajo las reglas que les han sido impuestas. Son, en cambio, agentes diestros que entienden y pueden modificar las condiciones institucionales en las que operan; tienen destrezas activas que les permiten —aunque sea de manera limitada— producir, reproducir y alterar su entorno social. Esto implica reconocer que el actual sistema de evaluación existe, en parte, gracias a la colaboración activa de los evaluados y, por lo tanto, es fundamental estudiar las condiciones y motivaciones de esa colaboración. Futuras investigaciones tendrán que incorporar esto como una parte de cualquier análisis y crítica de la evaluación del trabajo académico.

Referencias

- Aboites, 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.
- Aitkenhead, Decca (2013). Peter Higgs: I Wouldn't Be Productive Enough for Today's Academic System. *The Guardian*, 6 de diciembre.
- Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES) (2016). *Anuario de educación superior-licenciatura, ciclo 2015-2016*. www.anui.es.mx/iinformacion-y-servicios

- Bachelard, Gaston (2010). *La formación del espíritu científico*. México: Siglo XXI.
- Bartholomew, Robert E. (2014). Science for Sale: The Rise of Predatory Journals. *Journal of the Royal Society of Medicine*, 107 (10), 384-385.
- Bourdieu, Pierre (1995). *Respuestas. Por una antropología reflexiva*. México: Grijalbo.
- Broad, W. J. (1981). The Publishing Game: Getting More for Less. *Science*, 211 (4487), 13 de marzo, 1137-1139.
- Buendía, Angélica, Susana García Salord, Rocío Grediaga, Monique Landesman, Roberto Rodríguez-Gómez, Norma Rondero, Mario Rueda y Héctor Vera (2017). Queríamos evaluar y terminamos contando: alternativas para la evaluación del trabajo académico. *Perfiles Educativos*, 29 (157), 200-219.
- Chartier, Roger (1995). *Espacio público, crítica y desacralización en el siglo XVIII*. Barcelona: Gedisa.
- De Vries, Wietse y Germán Álvarez (2015). Can Reform Policies be Reformed? An Analysis of the Evaluation of Academics in Mexico. En Pavel Zgaga *et al.* (coords.), *Higher Education Reform: Looking Back-Looking Forward*. Fráncfort: Peter Lang Verlag, 257-272.
- De Vries, Wietse y Yadira Navarro (2013). Las controversias de la evaluación académica. *Avance y Perspectiva*, 5 (2), 1-10.
- Díaz Barriga, Ángel (2005). Los programas de estímulos del desempeño académico. En Ángel Díaz Barriga y Javier Mendoza (coords.), *Educación superior y Programa Nacional de Educación, 2001-2006. Aportes para una discusión*. México: ANUIES.
- Espeland, Wendy y Michael Sauder (2016). *Engines of Anxiety: Academic Rankings, Reputation, and Accountability*. Nueva York: Russell Sage Foundation.
- Espeland, Wendy y Michael Sauder (2007). Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology*, 113 (1), 1-40.
- Facultad de Ciencias Políticas y Sociales (FCPYS) (2008). *Instructivo para la dictaminación de concurso de oposición: ingreso de profesores de carrera*. México: UNAM.
- Frixione, Eugenio, Lourdes Ruiz-Zamarripa y Gerardo Hernández (2016). Assessing Individual Intellectual Output in Scientific Research: Mexico's National System for Evaluating Scholars Performance in the Humanities and the Behavioral Sciences. *PLoS ONE*, 11 (5). doi:10.1371/journal.pone.0155732
- Galaz Fontes, Jesús Francisco y Manuel Gil Antón (2013). The Impact of Merit-Pay Systems on the Work and Attitudes of Mexican Academics. *Higher Education*, 66 (3), 357-374.
- García de León, Antonio (2016). Tierra adentro, mar en fuera. Conferencia en el ciclo *La historia en los tiempos del cólera*. FCPYS, UNAM. 27 de octubre.
- García Salord, Susana (2013). Del recuento curricular a la evaluación diagnóstica y formativa. Un tránsito difícil pero necesario. *Perfiles Educativos*, 35 (núm. especial), 82-95.
- García Salord, Susana (2010). El curriculum vitae: entre perfiles deseados y trayectorias negadas. *Revista Iberoamericana de Educación Superior*, 1 (1), 103-119.
- García Salord, Susana, Rocío Grediaga y Monique Landesmann (2003). Los académicos en México hacia la constitución de un campo de conocimiento 1993-2002. En Patricia Ducoing (coord.), *Sujetos actores y procesos de formación*. México: Consejo Mexicano de Investigación Educativa, Colección Estados del Conocimiento, 115-268.

- Grediaga, Rocío (2006). Las políticas hacia los académicos en las últimas décadas. Cambios en la regulación de las trayectorias y el sistema de reconocimiento y recompensas de la profesión académica en México. *CPU-e, en Revista de Investigación Educativa*, 2, 1-72.
- Hicks, Diana y Paul Wouters (2015). The Leiden Manifesto for Research Metrics. *Nature*, 520, 429-431.
- Ibarra, Eduardo y Norma Rondero (2005). Regulación del trabajo académico y deshomologación salarial: balance general de sus ejes problemáticos. En T. Bertussi y G. González (coords.), *Anuario educativo mexicano: visión retrospectiva*. México: UPN-Miguel Ángel Porrúa-H. Congreso de la Unión, pp. 569-601.
- Instituto de Investigaciones Jurídicas (IIJ) (2014). *Contabilidad de obligaciones anuales mínimas*. México: UNAM.
- Jornet, Jesús (2016). La utilidad de la evaluación para mejorar la docencia. Conferencia. IISUE, UNAM. 18 de octubre.
- Mendoza Rojas, Javier (2011). *Financiamiento público de la Educación Superior en México: Fuentes de información y cifras del periodo 2000 a 2011*. Cuadernos de trabajo de la Dirección General de Evaluación Institucional, 6. México: UNAM.
- Ordorika, Imanol (coord) (2004). *La academia en jaque. Perspectivas políticas sobre la evaluación de la educación superior en México*. México: UNAM-Miguel Ángel Porrúa.
- Ordorika, Imanol y Roberto Rodríguez Gómez (2010). El ranking Times en el mercado del prestigio universitario. *Perfiles Educativos*, 32 (129), 8-22.
- Pérez Tamayo, Ruy (1996). "La pérdida de la inocencia del científico", *La Jornada*, 5 de agosto, p. 22.
- Rottenburg, Richard, Sally Merry, Sung-Joon Park y Johanna Mugler (coords.) (2015). *The World of Indicators: The Making of Governmental Knowledge through Quantification*. Cambridge: Cambridge University Press.
- Rueda, Mario (coord.) (2011). *¿Evaluar para controlar o para mejorar? Valoración del desempeño docente en las universidades*. México: UNAM.
- Secretaría de Educación Pública (SEP) (2017). "Lineamientos 2017 del Programa de Carrera Docentes (Fondo Extraordinario)". México: Secretaría de Educación Pública. <http://www.dgesu.ses.sep.gob.mx/Documentos/DSA%20go-bmx/lineamientos2017.pdf>
- Sindermann, Carl J. (1982). *Winning the Game Scientists Play*. Nueva York: Lenum Press.
- Sorokin, Pitirim (1956). "Quantophrenia", en *Fads & Foibles in Modern Sociology and Related Sciences*. Chicago: Henry Regnery Company, 102-173.
- Torres, Isaac (2016). Mejorar la evaluación de la calidad en el SNI podría hacerse de forma sencilla: Jorge Flores. *Foro Consultivo*, 9, 70-73.
- Universidad Autónoma Chapingo (UACH) (2013). *Reglamento del programa de estímulos al desempeño docente*. Texcoco UACH.
- Universidad Autónoma de Nuevo León (UANL) (2016). *Programa de estímulos. Ejercicio 2015-2016*. San Nicolás de los Garza: UANL.
- Universidad Autónoma Metropolitana (UAM) (2013). *Legislación universitaria*, México, UAM.
- Universidad Autónoma Metropolitana (UAM) (2016). *Informe de actividades 2015*. México: UAM.

- Vera, Héctor (2017a). El *homo academicus* y la máquina de sumar: profesores universitarios y la evaluación cuantitativa del mérito académico. *Perfiles Educativos*, 155, 87-106.
- Vera, Héctor (2017b). La evaluación cuantitativa del trabajo académico: tres analogías. *Sociológica*, 32 (90), 277-301.
- Vera, Héctor (2016). El sistema educativo y la ubicuidad de la evaluación. *Nexos*, 20 de enero. <http://educacion.nexos.com.mx/?p=132>.
- When, Francis (2006). *Marx's Das Kapital*. Nueva York: Atlantic Monthly Press.

Anexo 1

**Índice de siglas de las 27 instituciones en la Gráfica 1.
Entre paréntesis se indica el tipo o dependencia de cada institución**

UAM	Universidad Autónoma Metropolitana (Universidad pública federal)
Salud	Sistema Institucional de Investigadores en Ciencias Médicas (Secretaría de Salud)
IPN	Instituto Politécnico Nacional (Universidad pública federal)
CIDE	Centro de Investigación y Docencia Económicas (Centro Conacyt)
INAH	Instituto Nacional de Antropología e Historia (Secretaría de Cultura)
Instituto Mora	Instituto de Investigaciones Dr. José María Luis Mora (Centro Conacyt)
UdeG	Universidad de Guadalajara (UPES)
ITSON	Instituto Tecnológico de Sonora (UPES)
UV	Universidad Veracruzana (UPES)
UATX	Universidad Autónoma de Tlaxcala (UPES)
IJI, UNAM	Instituto de Investigaciones Jurídicas, UNAM (Universidad pública federal)
UPN	Universidad Pedagógica Nacional (Universidad pública federal)
UADY	Universidad Autónoma de Yucatán (UPES)
Colmich	Colegio de Michoacán (Centro Conacyt)
UQROO	Universidad de Quintana Roo (UPES)
Cinvestav	Centro de Investigación y de Estudios Avanzados (Universidad pública federal)
FCPYS	Facultad de Ciencias Políticas y Sociales, UNAM (Universidad pública federal)
BUAP	Benemérita Universidad Autónoma de Puebla (UPES)
Colmex	Colegio de México (Universidad pública federal)
UACAM	Universidad Autónoma de Campeche (UPES)
UACH	Universidad Autónoma de Chihuahua (UPES)
UGTO	Universidad de Guanajuato (UPES)
UASLP	Universidad Autónoma de San Luis Potosí (UPES)
UAS	Universidad Autónoma de Sinaloa (UPES)
UJED	Universidad Juárez del Estado de Durango (UPES)
UJT	Universidad Juárez Autónoma de Tabasco (UPES)
UAEM	Universidad Autónoma del Estado de Morelos (UPES)