

A LA CALIDAD POR THE QUANTITY (PORQUE LA CANTIDAD NO VALE). ALGUNAS REFLEXIONES SOBRE LOS CRITERIOS DE EVALUACIÓN DE LA CALIDAD DE LA INVESTIGACIÓN EN PSICOLOGÍA

Jesús Gil Roales-Nieto

Carmen Luciano

Universidad de Almería

RESUMEN

En este artículo se analizan los actuales sistemas de evaluación de la calidad investigadora en general y de las publicaciones en particular, con especial énfasis en los sistemas bibliométricos generados por el Institute for Scientific Information, discutiendo su aplicación a la ciencia de la psicología. Se analiza la importancia del idioma en el que se publica, la generalización que se hace del impacto de la revista al impacto de artículos específicos y las consecuencias que tiene todo ello en la determinación de la calidad de la producción científica en psicología. Finalmente, se discute la necesidad de cambio en los valores que controlan el sistema actual de evaluación y se presentan algunas propuestas de cambio.

Palabras clave: IMPACTO CIENTÍFICO, FACTOR DE IMPACTO, CALIDAD CIENTÍFICA, EVALUACIÓN INVESTIGADORA.

ABSTRACT

In this paper, we analyze the current evaluation system on the quality of the research, both in a general scope and then on publications. Especial emphasis is done about the bibliometric system generated by the Institute for Scientific Information, on the basis of its application to psychology as a social science. Some points are discuss: The impact of the language used in publicaciones; the function derived from the impact of the specific Journal to any of the specific pieces of research published in the context of the Journal; and, the consequences on both aspects in specifying the quality of the psychological scientific research. Finally, we discuss the need of a change in the values that controlled current evaluation system, and present some proposal to alter this state of affairs.

Key words: RESEARCH IMPACT, IMPACT FACTOR, RESEARCH QUALITY, RESEARCH EVALUATION.

Vaya por delante, y conviene decirlo cuanto antes y sin más prolegómenos, que quienes esto escriben han publicado y continuarán publicando en inglés y en castellano, y en revistas con impacto y sin impacto, participan de la puesta en marcha de una revista bilingüe, son miembros de comités editoriales de revistas nacionales e internacionales, y han pasado con éxito las evaluaciones pertinentes en el momento que les correspondió, confiando volver a pasarlas en el momento que les corresponda de nuevo, y están firmemente convencidos que la actividad científica debe y puede evaluarse.

Lo que se discute en este artículo, y probablemente subyace como mar de fondo a toda esta cuestión que ha llevado a dedicar una monografía al tema, es más los modos y los criterios. Es más la consideración del papel que debe jugar en el mundo nuestra cultura y nuestra ciencia. Bueno sería que el lector mantuviera siempre presente estas ideas iniciales mientras nos honra con la lectura de este escrito.

LA RAÍZ DE LAS COSAS O POR QUÉ LAS COSAS SON COMO SON

Empeñados en hacer de la Psicología una ciencia objetiva, fiscalista y "natural" podemos terminar por convertirla en una ciencia rígida, de cartón piedra, de constructos cosificados cuya medición y verificación se persigue por doquier.

Siendo así que la fiebre, adolescente y envidiosa, del objetivismo a ultranza ha terminado convirtiéndose en una compulsión cientifista (que no científica), se ha trasladado al terreno de lo social, de lo pragmático, de lo burocrático, de lo organizativo de la vida cotidiana del paisanaje científico de la psicología; al terreno de la valoración de lo que se hace (de lo que hacemos) como válido o inválido, bueno o malo, con calidad o mediocre, de un modo rigorista, maximalista y reduccionista, y en ocasiones arbitrario. Y es que si las ciencias serias, objetivas, naturales o "duras" (a las que se adora, se envidia y se trata de imitar sin medida y perspectiva), emplean el Science Citation Index (SCI) como tablas de la ley mosaico-científica (cuyo primer mandamiento es publicarás en inglés sobre todos las demás lenguas), no puede ser menos que la psicología, si quiere ser tenida por una ciencia a cualquier precio, termine por rendirse al Social Science Citation Index (SSCI), variante gentil de la ortodoxia bibliométrica, y acepte el factor de impacto que genera cada año como el equivalente a calidad.

Y una de las numerosas consecuencias que se han derivado de esta conversión a la religión del imperio científico-natural anglosajón -y de las más preocupantes- es la actitud exclusivista y excluyente, de que no puede existir más que Una Psicología "científica, internacional y de calidad", verdadera en esencia, y luz y guía del quehacer investigador, y el resto de psicologías se declaran no científicas, o no internacionales (locales, nacionales, paletas...) y, consecuentemente no de calidad (esto es, mediocres, limitadas, erróneas...); que a la vez debe ser Grande, esto es, mayoritaria y doctrina oficial, universal y universalizante (todos los hipotálamos son iguales y la neurociencia es su profeta); y, por supuesto, Libre, en la medida que la libertad del raciocinio científico sólo cabe expresarla "científicamente", a través de los acotados linderos de la ciencia natural, del método científico tomado del naturalismo, sin esencias culturales distorsionantes (la variabilidad como molestia).

Puede que estemos exagerando. Lo más probable es que lo estemos. Puede que estemos exacerbando el análisis (cosa que, a veces, es necesario como método para lograr la síntesis), pero también puede que la aparente exageración y exacerbación del análisis sea, en realidad, un mero disfraz del juicio anticipatorio, premonitorio, de dónde podemos acabar si no cambiamos el intento por establecer un modo único de pensar y proceder, y no enfrentamos la urgente y determinante tarea de construir socialmente una opción de psicología enraizada en nuestra cultura y nuestra lengua, aunque con visión integradora e instalada en la universalidad del comportamiento humano; dónde podemos acabar si no dejamos de plegarnos a los encantos del imperio angloparlante, y de las prácticas desintegradoras propias de la necesaria normalización y unicidad de pensamiento que requiere cualquier monopolio conceptual.

Visto así, y en un ejercicio de transliteración histórica, ¿sería justo dejar a Sigmund Freud sin sexenios si se le hubiera ocurrido retrasar setenta años su producción científica y presentarse a nuestras queridas evaluaciones, visto que publicaría en alemán y que probablemente lo haría en revistas que no entran en el sagrado SSCI? Encontraremos colegas que respondan afirmativamente a esta pregunta, y vean normal que Freud se quedara sin sexenios, porque lo suyo, de suyo, no es "ciencia" y sus publicaciones no cubren los estándares de calidad. Pero también encontraremos colegas que se lamenten de la falta de perspectiva de estos colegas de la justicia retrospectiva, entendiendo que ciertas aproximaciones en psicología no pueden ser tratadas como lo es la Astrología por la Astronomía. Ciertas ciencias resolvieron a lo largo de los siglos XV al XIX sus diferencias en cuanto a paradigma básico o universal, de manera que mientras que no hay más que una biología, una astronomía, una matemática y una física, para bien o para mal, aún contamos con más de una psicología, y quizá sea para siempre jamás, y quizá no sea tan malo como creímos que podía ser...

Y es que el fenómeno de la corriente principal o paradigma dominante puede que funcione en las disciplinas científicas tradicionales, al menos hasta cierto punto, pero la psicología, de momento, no es una ciencia de verdad o falsedad, o de paradigma central único e insalvable fuera del cual no se pueda hacer ciencia. De modo que

la corriente principal o dominante (que la hay, vaya si la hay) no puede confundirse con la correcta o verdadera, destinando al destierro científico a todas las demás por el mero hecho de ser minoritarias, declarándolas herejes, erróneas y no científicas.

La corriente principal, aunque esté dividida, a su vez, en múltiples corrientes, arroyos y torrenteras, significa en términos estadísticos que los psicólogos que piensan y escriben bajo su paraguas son mayoría frente a los psicólogos que piensan y escriben en otras maneras. Pero de ahí no puede deducirse que, de la misma forma que la inmensa mayoría de los médicos y biólogos participan de un modelo de pensamiento y acción en medicina y biología que se reconoce como paradigma científico exclusivo de ambas disciplinas (y todo lo demás es el desierto, o la selva, de lo paramédico, de la ciencia ficción, etc.), y dicha mayoría es una "garantía" de ciencia, en psicología podamos otorgar el mismo papel confirmador a la mayoría. En nuestra disciplina la opción mayoritaria no imprime carácter de soldado de la ciencia psicológica al investigador que forma parte de ella, ni otorga carácter exclusivo al paradigma en el que se sustenta dicha mayoría. La cantidad no garantiza la cualidad.

Ahora bien, el problema es cuando la cantidad acaba determinando la cualidad, porque se traduce en una mayor producción, un mayor número de revistas, un mayor número de entre-citas, etc. De tal manera que el impacto (medido asépticamente) de un artículo escrito de acuerdo al canon mayoritario tiene, de entrada, una probabilidad de ser citado, inmensamente mayor a la de cualquier artículo escrito bajo cualquiera de las herejías psicológicas minoritarias. Pero la calidad científica de ambos puede ser precisamente contraria a lo que dictan los números ofrecidos por el SSCI.

Por otro lado, vivimos en una sociedad tan complicada y competitiva que ya nada es posible en la viña del Señor sin que un, igualmente complicado y aparentemente riguroso, sistema de medida evalúe, clasifique y sentencie, priorizando o ninguneando, las mil y una opciones que se someten al sistema, sean éstas proyectos, profesores en busca de "sexenios" o grupos de investigación en demanda de unos euros por arrancar a los Planes de Investigación de cada una de nuestras universidades.

La cosa se agrava porque en España tenemos un sistema universitario y de investigación tan estatalizado, tan público él, tan dependiente de los euros de las numerosas administraciones que nos dirigen la vida, que la única solución posible es una burocratización a ultranza del sistema de distribución de fondos personales (la escasa pecunia adicional de los sexenios) o colectivos (las ayudas para investigar).

Como no podía ser de otra manera, la nueva Ley Orgánica de Universidades (LOU) ha venido a poner un poco de orden en un sistema demasiado voluntarioso y poco reglado y consciente de la complejidad que posee. Y así, la LOU instaura, reconoce e inaugura una cultura de la evaluación en toda regla, para que todos terminemos sabiendo, de una manera pública, objetiva e igualitaria, quienes "somos" cada uno y qué departamentos, centros y universidades ofrecen más y mejores servicios y rendimientos en investigación, más calidad, como es dado en decir ahora.

De hecho, la Ley establece en su artículo 32 la creación de la Agencia Nacional de Evaluación de la Calidad y Acreditación, y establece en el artículo anterior que "la promoción y la garantía de la calidad en las Universidades españolas, en el ámbito nacional e internacional, es un fin esencial de la política universitaria", señalando los cinco objetivos que persigue dicha garantía de la calidad. Hemos pasado pues, de la calidad como algo deseado, recomendado, recogido aquí y allí en forma subsidiaria o complementaria, a la calidad como uno de los fines esenciales del sistema universitario reconocido por Ley, a la calidad como algo necesario para sobrevivir. La cuestión de la evaluación de la calidad se convierte así en un aspecto central de debate para el futuro del sistema universitario español. Obviamente, dadas las características de este número monográfico y el compromiso que alcanzamos como autores de este artículo, circunscribiremos nuestra humilde aportación al ámbito estricto de la calidad en la producción científica y de su correspondiente evaluación.

Y como hemos comprobado que se avecina un tiempo en el que vamos a ser concienzudamente evaluados queramos o no, y en todos los planos, bueno es que reflexionemos sobre qué medir y cómo hacerlo, para que dicha evaluación no termine convirtiéndose

en la actual devaluación selectiva de cierto tipo de ciencia en la que ha terminado por convertirse el intento uniformador de establecer que la calidad pasa por New York y no por Barcelona (por citar dos ejemplos al azar), y que como suele decir con una gran dosis de socarronería Vicente Pelechano, el idioma de Cervantes "debe ser empleado no sólo para hacer novelas y poesía, sino también la mejor ciencia posible".

Y es que tal parece que los psicólogos (al menos, ciertos psicólogos) se hayan vuelto más "impactistas" que los propios creadores del impacto, surgido principalmente para ciencias distintas de la nuestra, pero que ha terminado por invadirlas casi todas. Pero también conviene dejar claro que la puesta en cuestión del sistema que tratamos no afecta en exclusiva a la psicología o a las ciencias cercanas. No se trata de una revuelta local contra un sistema incontestado en el resto de la disciplinas. Quien mantenga esa idea conviene que considere la existencia de críticas duras y contundentes que denuncian la relatividad, sesgo y consecuencias adversas que está produciendo el culto ciego al factor de impacto de las revistas como medida de la calidad de las publicaciones. Y estas críticas se encuentran en disciplinas tan variadas como la medicina (p.ej., Camí, 1997; Fontanarosa, 1998; Gastel, 2001; Gonzalo Sanz, 1995; Hecht, Hecht y Sandberg, 1998; Reyes, 1998; Van Diest, Holzel, Burnett y Crocker, 2001), la ecología (p.ej., Altmann y Gorman, 1998; Carrascal, 1997; Kokko y Sutherland, 1999), las ciencias biomédicas (p.ej., Coleman, 1999), la zoología (p.ej., Altmann y Gorman, 1999), las ciencias sociales (p.ej., Fur, 1995), la farmacología clínica (p.ej., Rostami-Hodjegan y Tueker, 2001), las neurociencias (Gillis, 1999), la bioquímica (p.ej., Seglen, 1989) o la psiquiatría (Elkis, 1999). Y estas críticas provienen de contextos científicos hispanos, anglosajones y europeos, de manera que son tan internacionales como la pretendida "internacionalidad" del factor de impacto.

Es el momento, pues, de reflexionar sobre cómo definir lo que entendemos por publicaciones de calidad y con impacto, y cómo establecer ambas cosas. No vaya que, por no saber medir, acabemos perdiendo definitivamente el sentido de la medida. Esto es, acabemos pontificando y elevando a la categoría de dogma irrefu-

table que la calidad en las publicaciones de psicología significa exclusivamente publicar en el idioma de William Shakespeare y el pato Donald, y en las revistas que aparezcan en las listas del Journal of Citation Reports (JCR), del negocio privado de la empresa norteamericana radicada en Philadelphia llamada *Institute for Scientific Information (ISI)*.

ESTADO GENERAL DEL ASUNTO: EL USO DE FACTORES DE IMPACTO, LA CUESTIÓN DEL IDIOMA Y EL SOMETIMIENTO AL RÉGIMEN ISISTA

El uso de factores de impacto como criterio central de calidad

Para decirlo pronto y de una vez, el factor de impacto (FI) de un artículo como réplica directa del FI de la revista en la que se publicó es una impostura. El factor de impacto que el JCR publica, lo es de la revista en la que se publicó el artículo (en concreto de sus dos últimos años), pero no de ninguno de sus artículos en sí y tomados individualmente. Un artículo publicado en una revista con factor de impacto de 23,5 no tiene necesariamente que haber sido leído por 23,5 veces más personas que un artículo publicado en una revista con factor de impacto 1. Puede que un artículo en concreto, aún habiendo sido publicado en una revista de alto impacto, haya sido, en realidad, leído y citado sólo por el grupo de colegas que se dedican al mismo tema de investigación y que pueden no pasar de algunas docenas entre los miles de lectores de la revista, mientras que otro artículo de una revista de escaso impacto puede haber sido leído (y después citado) por centenares de lectores de esta revista y otras revistas ninguna de las cuales se encuentra en el SSCI (*Social Science Citation Index*) aunque cumplan con sus criterios formales y lleven a cabo serios procesos de selección. Si medimos la calidad por el impacto, ¿qué significa realmente impacto?

El concepto de impacto ha sido bien precisado por Camí (1997) cuando lo define como "una cualidad que vendría a interpretar la supuesta repercusión inmediata de una investigación en un sistema de comunicación científica imperfecto". Es, en esencia, más un

indicador de aparente difusión o aparente visibilidad entre la comunidad científica de moda o dominante. Más cercano a la fama que a otra cosa, aunque la fama no es ni buena ni mala sino todo lo contrario.

Como señalamos anteriormente, el impacto sale de las citas y las citas no siempre salen de donde deben salir. Es bien conocido que la distribución de las citas recibidas por los artículos publicados en una misma revista se distribuyen asimétricamente, fenómeno que ejemplificó Seglen (1992) cuando comprobó que el FI de una revista (en concreto, *The Biochemical Journal*) bajaba de 6,2 a 3,5 cuando se excluían del análisis los tres artículos más citados.

El mismo autor señaló que, en casi todas las revistas, alrededor del 15% de los artículos acaparan más del 50% de todas las citas, lo que indicaría que en todas las revistas, independientemente de su FI, existe un contingente de artículos "gregarios" o de "pelotón" a la hora de calcular el FI, pero que, cuando se generaliza del todo a las partes, participan de la fama de los ganadores en el *Tour* de las citas. Está comprobado hasta la saciedad que el FI de una revista no garantiza a priori que, por aparecer en ella, un artículo vaya a recibir el promedio de citas que dio lugar a dicho FI.

Y es que hasta Eugene Garfield (1996) -recuérdese, el fundador en los primeros años 60 del siglo pasado del factor de impacto junto a Irving H. Sher- ha señalado la curiosa anécdota (válida como método de reducción al absurdo) del artículo sobre el método para la determinación de proteínas que Lowry publicó y que ha acumulado el récord de 250.000 citas desde 1952, aún y cuando el propio autor nunca lo consideró como el mejor de sus trabajos publicados. De donde se deduce que el poder otorgador de calidad del FI es tal que se impone incluso al juicio de los propios autores sobre sus propios trabajos (¡oiga, mire, lo suyo es de calidad aunque usted no lo crea!).

Los gurús burocratizados de la ciencia psicológica (quizá recelosos del *status* de las ciencias tradicionales) si substituyen el juicio histórico (el único que mide el verdadero impacto a través del tiempo de lo investigado y escrito), por el juicio estadístico del impacto torticeramente obtenido, acabarán propiciando el encorsetamiento de la ciencia psicológica y de la libertad de pensamiento y acción, al pretender establecer, apriorísticamente, qué idea, hallazgo o

conjetura tiene o no calidad por haber sido o no publicada aquí o allá, sin esperar a que sea el propio, reposado y certero juicio de la historia el que determine si tal artículo publicado en la más impactante revista de la iglesia internacional-ISIsta, en realidad, no ha supuesto el más mínimo avance en su disciplina.

Otro fenómeno que el ISI no analiza al elaborar sus factores de impacto es el efecto de la red de citaciones. Esto es, unos cuantos equipos de científicos de dos, tres o cuatro países con un tema común, pueden conseguir un considerable número de publicaciones en dos o tres revistas cada dos o tres años; si se citan sistemáticamente entre ellos elevarán notoriamente el factor de impacto de las citadas revistas, aunque puedan no haber conseguido ni una sola cita "externa" a la red, o peor aún, aunque sólo hayan conseguido citas "externas" de autores que hayan puesto de manifiesto la escasa valía de los trabajos publicados (citas son citas, y el ISI no entra en estas menudencias)

La cuestión del idioma como elemento supra o minusvalorador

Defender que la calidad deviene de la publicación en inglés y en revistas "internacionales", parece obra de reprimidos culturales que se hayan criado aprendiendo de sus mayores que "todo lo bueno viene de fuera y que lo nuestro es necesariamente peor", sin percatarse que el engaño va siempre más lejos que las sospechas como bien dijo La Rochefoucauld, y que el autoengaño va siempre muchísimo más lejos que las autosospechas, como nos permitimos añadir nosotros. Un ejemplo real de este complejo aparece reflejado en el cuadro 1, que recoge los criterios de puntuación de publicaciones de calidad elaborados por un juez único, omnipotente e infalible, para la puntuación de las publicaciones de calidad de los grupos de investigación de psicología en una convocatoria de ayudas del Plan Propio de Investigación de una universidad española. Obsérvense dos cosas. Primera, como es que el idioma determina una puntuación diferencial para revistas con las mismas condiciones de impacto: un artículo en una revista internacional especializada con revisores, que aparezca en el *Citation Index*, con un índice de impacto

Cuadro 1.- Criterios para la evaluación de las publicaciones de calidad de Psicología elaborados para un Plan Propio de Investigación de una universidad española (es transcripción literal, se ha respetado fielmente lo escrito en el original excepto en el énfasis –subrayado–, que es nuestro)

La evaluación de la calidad de un trabajo debe venir determinada por su impacto en el conjunto de la comunidad científica a la que se dirige. Una forma de determinar dicho impacto de forma objetiva, y ajena a las excelencias del trabajo que los propios autores pueda realizar de forma subjetiva sobre el mismo, es atender a la relevancia local, nacional o internacional de la revista/editorial en la que ha sido publicado. Considero que un trabajo publicado en un medio de ámbito local difícilmente podría entrar a considerarse dentro de un apartado que trata de discriminar aquellos trabajos con un impacto real y objetivo como es el de Publicación de Calidad, ya que su impacto, independientemente de la calidad del trabajo, y los requerimientos para su publicación, son mínimos. Por tanto, como criterio general, a todas aquellas publicaciones de ámbito local se les asignará la puntuación 0. Las publicaciones de ámbito nacional e internacional se evaluarán con los criterios que se especifican a continuación, siempre teniendo en cuenta, en el caso de artículos, su aparición en el Citation Index, único índice generalmente aceptado por la comunidad científica y otros órganos de evaluación (ej. ANEP) como indicador del impacto de las mismas.

Artículos en revistas científicas especializadas

Artículos en revistas internacionales especializadas con revisores, que aparezcan en el Citation Index:

Con índice de impacto por encima de la mediana 3
Con índice de impacto por debajo de la mediana 2

Artículos en revistas nacionales especializadas con revisores, que aparezcan en el Citation Index:

Con índice de impacto por encima de la mediana 2
Con índice de impacto por debajo de la mediana 1

Artículos en revistas nacionales que no aparezcan en el *Citation Index*..... 0

Acogerse al régimen ISIsta como única medida de calidad tiene, además, consecuencias nefastas para el desarrollo de nuevas publicaciones. Tal como Camí (1997) ha sentenciado certeramente "la mezcla y/o la comparación de trabajos de distintas disciplinas basadas únicamente en el FI de la revista es una solemne barbaridad" (pág., 519).

Estamos convirtiendo la sana actividad de publicar en una cuestión comercial que se rige por las leyes del *marketing*, de manera que se atiende prioritariamente a las leyes del mercado. Así, se busca "colocar" un artículo en determinada revista porque ello "garantiza" la imagen de calidad del artículo e incrementa su valoración en la comunidad científica ISIsta. De manera que, un ejercicio singular puede ser la huida hacia el lenguaje dominante y hacia la terminología al uso, de tal forma que los científicos de la psicología terminan "acomodando" su producto al mercado en el que lo quieren vender; aunque ello obligue a utilizar estadística cuando no es necesaria o superflua, o a suprimir información que los capataces del mercado consideran innecesaria o irrelevante, o a cambiar los conceptos, o a acomodar los datos a la teoría de moda, o a lo que hubiera menester. Justificando todo ello con pensamientos simples, flexiones más que reflexiones, que lo señalan como el único modo de lograr el éxito en un sistema que lo define como la colocación de artículos en las revistas de impacto.

Proceso de moldeamiento del lenguaje y los modos científicos de acuerdo a la corriente dominante, que termina interiorizándose como regla de conducta adaptativa y eficaz para el científico exitoso gracias a las contingencias provistas por el sistema. El éxito es la adaptación al sistema. Todo ello significa que existe una configuración apriorística que determina el valor intrínseco de una publicación en base a su utilidad para superar los filtros del sistema de publicación.

LA SESUDA CUESTIÓN DE LA EVALUACIÓN DE LOS "SEXENIOS"

La evaluación de los sexenios sigue, en buena medida aunque dependiendo mucho de los evaluadores de turno, las reglas del

marketing, convirtiendo en bueno todo lo que vende mucho (todo lo que tiene "audiencia" tomada ésta como factor de impacto de la revista en la que se publica) y minusvalorando todo lo que, al quedar fuera del circuito ISIsta de las tendencias hegemónicas, no "tiene audiencia".

De esta manera, vale más un artículo hábilmente colocado en una revista "impactante" aunque no lo lea nadie, que dos ediciones de un buen ensayo, o un voluminoso tratado que revise, reflexione y replantee alguna cuestión fundamental. Es el caso, por ejemplo, de *El descubrimiento del inconsciente* (Ellenberger, 1970), libro que puede que le reportase menos puntuación a su autor (si tuviera que solicitar los tales sexenios), que un artículo sobre la diferencia en percepción de una gama de colores en esquizofrénicos *versus* no esquizofrénicos, publicado en alguna de las numerosas revistas del imperio dominante. Aunque para hacer lo primero haya que trabajar diez años y para lo segundo quizá baste con tres meses y un puñado de ordenadores frente a los que colocar a unas docenas de pacientes y otras de estudiantes de psicología "voluntarios".

Por otro lado se encuentra la puesta en escena de la evaluación. Afortunadamente, la nueva ley obligará a la reforma profunda del sistema actual con la creación de la Agencia Nacional de Evaluación y Acreditación. Pero en las circunstancias actuales, es completamente inadecuado que un solo juez, sea el juez que sea y aunque técnicamente se ampare en una decisión administrativamente tomada por un "órgano colegiado" como es la comisión o subcomisión correspondientes, sea el que determine si un investigador cumple o no los criterios establecidos para la evaluación positiva, y todo ello sin acceso directo a la producción científica del evaluado. Cuando ni tan siquiera un simple artículo es evaluado por un solo revisor en cualquier revista que se precie, choca que algo cuyas consecuencias son mayores, y en la que un falso negativo inhabilita para un nuevo intento hasta pasados tres años, sea dejado en manos de un solo juez para el caso de la psicología: una disciplina con 6 dispares áreas de conocimiento que obligaría al juez único a unos conocimientos enciclopédicos impropios de los tiempos que corren en la ciencia.

Pero es que, además, y ello es doblemente grave porque impide conocer las corruptelas que puedan darse, es imposible detectar los

por encima de la mediana recibirá 3 puntos, que serán 2 si la revista es nacional, aunque su impacto también esté por encima de la mediana. Puede que el soberbio científico autor de estos criterios desconozca que, de aplicarse sus normas en USA, sus revistas internacionales pasarían a ser nacionales. Segunda, la dramática confesión que contienen las líneas centrales (desde "Considero..." hasta "...son mínimos") en las que queda meridianamente claro que la evaluación de las publicaciones de CALIDAD se hace por el impacto e INDEPENDIENTEMENTE DE LA CALIDAD DEL TRABAJO. Es uno de los más claros ejemplos de impactolatría que cabe esperar.

Por este complejo de internacionalidad, que obliga a quien lo padece a escribir en la lengua del imperio para lograr el certificado de existencia científica, a modo de auto habeas corpus afirmativo, se niega a los psicólogos y otros profesionales de habla hispana, leer artículos de calidad en su lengua materna, debiendo hacerlo necesariamente en la de Bill Clinton. Y, con ello, se cae en el error puesto de manifiesto por Gastel (2001), cuando denunciaba que escribir en inglés para científicos de otra lengua es una forma de escamotear a su contexto cultural el valor de sus hallazgos, olvidando que su impacto real, por social y "práxico", sería probablemente mayor si sus mejores hallazgos los sirvieran en su lengua materna, dando acceso a los mismos a toda la comunidad científica de su país de origen. En algo de esto pensaba el editor científico del *Neederlands Tijdschrift voor Geneeskunde* (Revista de Medicina de los Países Bajos), cuando señalaba que "la investigación de alcance nacional, con consecuencias nacionales, y financiada con fondos públicos debería ser publicada en la lengua nacional".

Otra posibilidad es exigir a todos los ciudadanos de este país (y de cualquier otro) el conocimiento hablado y escrito de la lengua inglesa, de modo que sus élites profesionales no se vean privadas de la repercusión de los hallazgos más brillantes de nuestros investigadores en el ámbito de la psicología, no vaya a ser que un extraordinario método de tratamiento del tabaquismo, la depresión o el maltrato doméstico, pase desapercibido, se malentienda o tarde años en conocerse porque los investigadores necesitaron colocar su artículo en una revista "internacional" escrita en inglés, editada en Wisconsin e impresa en Singapur, pero con un FI de 10.

Porque, claro, como lo del impacto es algo cuantitativo (de número) y los psicólogos angloparlantes son más, y los que hablan o leen en español casi inexistentes, publicar en español, por ejemplo en esta revista, supone, de entrada, renunciar al logro de no más que un modesto FI, que es lo que lograrán, o han logrado ya, las mejores revistas en nuestra lengua, porque el mercado no da para más.

El sometimiento al régimen "ISIsta" como único referente de calidad

De esta manera, la miopía lleva a considerar el régimen "ISIsta" (esto es, la aplicación del FI de las revistas para valorar la calidad de los artículos) como el mejor, si no el único, referente de calidad para la evaluación de la actividad investigadora en la psicología hispanohablante.

El régimen ISIsta se está empleando por doquier como casi única manera de establecer la calidad de la investigación, establecer prioridades en la financiación y evaluar el resultado del trabajo de personas, grupos y universidades. Sin embargo, tal como se está utilizando, de manera presuntamente aséptica y general, sin tener en cuenta ni siquiera las limitaciones que el propio ISI y el creador del factor de impacto han señalado (p.ej., Garfield, 1979, 1996, 2001; Garfield y Welljams-Dorof, 1992). Se trata, por tanto, de un falso sistema de calidad y las razones han sido expuestas con clarividencia y detalle en numerosas ocasiones (p.ej., Brunstein, 2000; Buela Casal, 2001; Coleman, 1999; Elkis, 1999; Fava y Ottolini, 2000; Furr, 1995; Gastel, 2001; Gillis, 1999; Gisvold, 1999; Gómez y Bordons, 1996; Gonzalo Sanz, 1995; Harder, 2000; Hemmingsson, Edgren, Mygind, y Skjennald 2002; Hemmingsson, Mygind, Skjennald y Edgren, 2002; Krell, 2002; Moed, 2002; Moed y Van-Leeuwen, 1995; Moed, Van-Leeuwen y Reedijk, 1998; Reyes, 1998; Rodríguez, Rodríguez y Matesanz, 2000; Rostami-Hodjegan y Tucker, 2001; Rousseau y Spinak, 1996; Seglen, 1989, 1997a,b, 1998; Smith, 1998; Takahashi, Aw y Koh, 1999; West, 1996).

Un sistema de autoperpetuación y endogámico (tanto criticar la endogamia universitaria para acabar elevando a los altares a un sistema ferozmente endogámico, es una desagradable paradoja).

disciplinas), humanos que viven en sociedad y participan de una herramienta común de interacción que llamamos lenguaje. Y los objetivos de una ciencia especialmente social, recaen en la explicación de los fenómenos que estudia, mas ni sólo ni tanto en sí mismos ni por sí mismos, sino que sus objetivos se extienden necesariamente a la predicción y control de los fenómenos susceptibles de explicación.

Así pues, la calidad científica en las disciplinas básicas y aplicadas como la nuestra, debe incluir necesariamente un aspecto pragmático, un aspecto que contemple la utilidad de lo investigado y publicado (p.ej., Hayes, 1993; Smith, 2001), no sólo para incrementar la leyenda o realidad sobre la calidad de un investigador o un equipo de investigadores, sino también la utilidad para los objetivos primarios por los que se investiga: disponer de conocimiento que nos permita solucionar problemas humanos y sociales.

Un principio irrenunciable es el que establece que la calidad de una publicación es algo intrínseco a la publicación y no dependiente del continente en el cual se sirva. No depende, por tanto, de factores externos a la misma como pueden ser el acierto o error en la elección de la revista en la que verá la luz, o el idioma en el que se sirva a los lectores. Un estudio o trabajo teórico tiene o no calidad por sí mismo y sea o no conocido (calidad y fama no son relaciones necesariamente contingentes), sea o no leído. Confundir calidad con publicidad o fama es un error de principio en la evaluación de los productos de la actividad investigadora. El impacto de un procedimiento terapéutico sobre la práctica profesional de miles de psicólogos puede verse sólo moderadamente reflejado en el impacto estadístico del promedio de citas, pero su influencia o afectación sobre el fenómeno que estudió puede haber supuesto mucho más, en términos de avance de su disciplina, que un artículo profusamente citado. Medir esto es complicado, pero reducir el asunto a meros términos de número es mucho peor.

Un trabajo de calidad es un trabajo bien hecho y que logra los objetivos que se propuso. Es un trabajo que aporta conocimiento y soluciones a problemas planteados, que ha sido hecho con corrección y rigor metodológico, con originalidad y eficacia en el logro de los objetivos planeados. De la misma manera que el refranero nos

enseña que "la mona, mona es y mona se queda aunque se vista de seda", deberíamos saber que la calidad, calidad es y calidad se queda aunque no se vista siempre de *silk*.

Calidad, relevancia o presencia e impacto, son tres conceptos diferentes para propósitos diferentes y dependientes de variables diferentes, que pueden o no coincidir en el espacio y el tiempo, que pueden confluír o alejarse dramáticamente en según qué casos y según qué contextos.

Otra posibilidad, que parece irrenunciable, es la elaboración de sistemas "articulados" de medida, esto es, basados en el artículo no en la revista. Si la cualificación se otorga al artículo lo lógico es medir el impacto del artículo, no de la revista. De esa manera se impide el actual sistema de falsa calidad que se está imponiendo en psicología, de tal forma que investigaciones mediocres, que simplemente replican directa o sistemáticamente lo que hacen sus referentes en la metrópoli, pueden convertirse en publicaciones de "calidad" por el simple hecho de haber sido aceptadas en revistas con impacto, al tratar temas de moda y estar en línea con "lo que se hace en ese momento". Es decir, necesitamos un sistema articulado de medida de la calidad que permita valorar el impacto en cuanto repercusión, de un específico artículo sobre otros trabajos de investigación o sobre la diseminación de técnicas y procedimientos útiles para su aplicación en la solución de problemas personales o sociales. Un sistema articulado no contradice el culto al impacto, simplemente lo verifica y permite prescindir del canon ISIsta recuperando la dignidad científica de la ciencia en español.

Hacer esto es más difícil, pero si resulta tan difícil como para no merecer la pena el coste que supone, no convirtamos en doctrina el resultado de una economía de gestos: como es más fácil y barato medir el impacto de las revistas entre las revistas de las que se dispone gratuitamente (lo que hace el *ISI*), establezco que los artículos de calidad son "aquellos publicados en revistas de calidad" y se entiende por tal lo más arriba analizado. Hacer esto es algo muy parecido a un fraude de ley.

Es un error convertir una evaluación cuantitativa en un resultado cualitativo que tendrá trascendencia social, económica y personal. Si numerosos análisis han dejado claro que tomar como principal

falsos positivos, en la medida que el resultado de la evaluación no es público, somos un gran país pequeño en el que casi todos nos conocemos y nadie puede recurrir la ocurrencia de un informe positivo prevaricador si tal cosa se produjese.

ALTERNATIVAS

Como muchas cosas en este mundo, la evaluación de los méritos investigadores y de la calidad de la producción científica requiere profundos y amplios cambios. Contamos con la experiencia suficiente como para llegar a la conclusión de la necesidad de una evaluación cuanto menos múltiple, que no única, y cuanto menos diversa, que no sólo cuantitativa. No es nuestra intención ofrecer un despliegue de alternativas claras y concisas. Primero, porque es más clara la necesidad de cambio que los cambios necesarios (cuya elaboración requiere de discusión, justificación, puesta a prueba y consenso). Segundo, porque ello corresponde a quienes tienen la responsabilidad en sus manos. Por tanto, entiéndase lo que sigue como meras reflexiones que puedan aportar ideas para el sano y necesario debate que, por iniciativa del director de esta revista, estamos llevando a cabo en este número especial.

Parece claro, tras lo expuesto aquí y en muchos otros lugares, que debemos luchar contra lo que ha sido definido desde la medicina como una enfermedad que está afectando a la comunidad científica, que algunos han bautizado como impactología (Camí, 1997) y otros como impactitis (Van Diest, Holzel, Burnett, Crocker, 2001), y que nosotros preferimos considerar como una nueva forma de idolatría con su liturgia, sus misterios, sus sacramentos, sus templos y su sacerdocio piramidalmente organizado en una jerarquía "científicamente establecida". Frente a esto, algunas ideas (que se recogen en el cuadro 2) deben permanecer claras y manifiestamente incorporadas a nuestros valores relacionados con la evaluación de nuestra actividad investigadora.

No podemos olvidar que la psicología es una ciencia especialmente social, porque trata de los humanos como un todo (las partes del todo ya son exhaustivamente estudiadas por un ejército de

Cuadro 2.- Consideraciones a tener en cuenta en la evaluación de las publicaciones como medida de calidad investigadora

1. Publicar no es sinónimo de contribución a la ciencia, por lo que la evaluación de la calidad de las publicaciones es necesaria. También se hace necesario definir qué se entiende por actividad investigadora.
2. Debe distinguirse claramente entre calidad, relevancia e impacto. De una evaluación meramente cuantitativa no pueden derivarse resultados cualitativos de calidad/no calidad.
3. El impacto de los artículos publicados en una revista NO es el impacto de la revista, por lo que éste no puede generalizarse a artículos específicos.
4. La evaluación de la calidad de una publicación concreta debe combinar métodos objetivos y subjetivos, debe hacerse con perspectiva y debe depender de la publicación en sí misma, no del continente en el que ésta ha sido servida.

- Carrascal, LM.** (1997). La referencia bibliográfica como medida de "utilidad científica". *EtoloGüía*, 15: 17-30.
- Coleman, R.** (1999). Impact factors: use and abuse in biomedical research. *The Anatomical Record*, 257: 54-57.
- Elkis, H.** (1999). Fatores de impacto de publicações psiquiátricas e produtividade científica. *Revista Brasileira de Psiquiatria*, 21: 231-236.
- Ellenberger, HF.** (1970) *The discovery of the unconscious*. New York: Basic Books. (Versión española en Ed. Gredos: Madrid, 1976).
- Fava, GA. y Ottolini, F.** (2000). Impact factors versus actual citations. *Psychotherapy and psychosomatics*, 69: 285-286.
- Fontanarosa, PB.** (1998). Impact factors and emergency medicine journals. *Annals of Emergency Medicine*, 31: 107-109.
- Furr, A.** (1995). The relative influence of social work journals: impact factors vs. core influence. *Journal of Social Work Education*, 31: 38-45.
- Gastel, B.** (2001). Assessing the impact of investigators' work: beyond impact factors. *Canadian Journal of Anesthesia*, 48: 941-945.
- Gillis, JM.** (1999). Overuse of impact factors suppresses controversial ideas. *Nature Neuroscience*, 2: 101.
- Gisvold, SE.** (1999). Citation analysis and journal impact factors -is the tail wagging the dog?. *Acta Anaesthesiologica Scandinavica*, 43: 971-973.
- Garfield, E.** (1979) Is citation analysis a legitimate evaluation tool? *Scientometrics*, 1: 359-375.
- Garfield, E.** (1996). How can impact factors be improved?. *British Medical Journal*, 313: 411-413.
- Garfield, E.** (2001). Impact factors, and why they won't go away. *Nature*, 411: 522.
- Garfield, E. y Welljams-Dorof, A.** (1992) Citation data: their use as quantitative indicators for science and technology evaluation and policy-making. *Science and Public Policy*, 19: 321-327.
- Gómez, I. y Bordons, M.** (1996) Limitaciones en el uso de los indicadores bibliométricos para la evaluación científica. *Política Científica*, 46: 21-26.
- Gonzalo Sanz, LM.** (1995). Índice de impacto: pros y contras. *Revista de Medicina de la Universidad de Navarra*, 39: 4-5.
- Harder, DR.** (2000). Impact factors and the competitive nature of journal publishing. *American Journal of Physiology. Heart and Circulatory Physiology*, 279: H457.
- Hayes, SC.** (1993). Goals and the varieties of scientific contextualism. En SC. Hayes, LJ. Hayes, TR. Sarbin y HW. Reese (Eds.), *The varieties of scientific contextualism* (págs., 11-27). Reno, Nv.: Context Press.

- Hecht, F., Hecht, BK. y Sandberg, AA.** (1998) The journal "impact factor": a misnamed, misleading, misuse measure. *Cancer Genet. Cytogenet.*, 104: 77-81.
- Hemmingsson, A., Edgren, J., Mygind, T. y Skjennald, A.** (2002). Use of impact factors. *Lancet*, 359: 153.
- Hemmingsson, A., Mygind, T., Skjennald A. y Edgren, J.** (2002). Manipulation of impact factors by editors of scientific journals. *American Journal of Roentgenology*, 178: 767.
- Kokko, H. y Sutherland, W.J.** (1999). What do impact factors tell us?. *Tree*, 14: 382-384.
- Krell, FT.** (2002). Why impact factors don't work for taxonomy. *Nature*, 415: 957.
- Moed, HF.**(2002). The impact-factors debate: the ISI's uses and limits. *Nature*, 415: 731-732.
- Moed, HF, Van-Leeuwen, ThN** (1995) Improving the accuracy of the Institute for Scientific Information's Journal Impact Factors. *Journal of the American Society for Information Science*, 46, 461-467.
- Moed, HF, Van-Leeuwen, ThN y Reedijk, J.** (1998) A new classification system to describe the ageing of scientific journals and their impact factors. *Journal of Documentation*, 54, 387-419.
- Reyes, H.** (1998). El "factor de impacto" y el impacto de las revistas médicas. *Revista Médica de Chile*, 126: 135-138.
- Rodríguez, A., Rodríguez, JC. y Matesanz, R.** (2000). La publicación científica y su lectura crítica. Normas para la publicación de un artículo científico. Factor impacto. *Nefrología*, 20, Suplemento 6: 29-33.
- Rostami-Hodjegan, A. y Tucker, GT.** (2001). Journal impact factors: a "bioequivalence" issue? *British Journal of Clinical Pharmacology*, 51: 111-117.
- Rousseau, R. y Spinak, E.** (1996) Do a field list of internationally visible journals and their journal impact factors depend on the initial set of journals? A research proposal. *Journal of Documentation*, 52, 449-456.
- Seglen, PO.** (1989). From bad to worse: evaluation by Journal Impact. *Trends in Biochemical Science*, 14: 326-327.
- Seglen, PO.** (1992) How representative is the journal impact factor? *Research Evaluation*, 2: 143-149.
- Seglen, PO.** (1997a) Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314, 498-502.
- Seglen, PO.** (1997b). Citations and journal impact factors: questionable indicators of research quality. *Allergy*, 52: 1050-1052.
- Seglen, PO.** (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthopaedica Scandinavica*, 69: 224-229.

- Smith, R.** (1998) Unscientific practice flourishes in science. *British Medical Journal*, 316, 1036-1040.
- Smith, R.** (2001) Measuring the social impact of research. Difficult but necessary. *British Medical Journal*, 323, 528-529.
- Takahashi, K., Aw TC. y Koh, D.** (1999). An alternative to journal-based impact factors. *Occupational Medicine*, 49: 57-59.
- Van Diest, P.J., Holzel, H., Burnett, D. y Crocker, J.** (2001). Impactitis: new cures for an old disease. *Journal of Clinical Pathology*, 54: 817-819.
- West, R.** (1996). Impact factors need to be improved. *British Medical Journal*, 313: 1400.