

Manipular Google Scholar Citations y Google Scholar Metrics: simple, sencillo y tentador

Emilio Delgado López-Cózar¹, Nicolás Robinson-García¹ y Daniel Torres-Salinas²
EC3: Evaluación de la Ciencia y de la Comunicación Científica

¹Universidad de Granada

²Universidad de Navarra

edelgado@ugr.es; elrobin@ugr.es; torressalinas@gmail.com

RESUMEN

La aparición de Google Scholar Citations y Google Scholar Metrics puede suponer una auténtica revolución en el mundo de la evaluación científica porque pone al alcance de todos los científicos herramientas para la medición bibliométrica. Con el fin de alertar a la comunidad científica de la facilidad que existe para manipular los datos e indicadores bibliométricos que proporcionan estos productos de Google, se ha realizado un experimento consistente en la manipulación del impacto de los miembros de un grupo de investigación y de sus trabajos tal como son reflejados en los perfiles de autores que ofrece Google Citations así como del número de citas e índice h de una muestra de revistas. Para ello se crean seis documentos producidos por un autor falso que emiten citas a todos los trabajos producidos por los tres autores y se suben a la página personal de un investigador de la Universidad de Granada. El resultado del experimento ha supuesto un incremento de 774 citas en los 129 trabajos producidos (con 6 citas acumuladas a cada trabajo) que revertido un incremento del índice h de los autores y de las revistas en los que se publicaron los trabajos citados. Se analiza el efecto perverso que este tipo de prácticas puede ocasionar en Google Citations y Google Scholar Metrics. Finalmente se ofrecen unas reflexiones finales sobre los efectos que pueden tener este tipo de prácticas y la falta de sistemas de control que ofrecen estos productos

PALABRAS CLAVE

Google Citations / Google Scholar Metrics / Revistas / Fraude científico / Dishonestidad científica / Manipulación / Citas / indicadores bibliométricos / H Index / Evaluación / investigadores / autores

Manipulating Google Scholar Citations and Google Scholar Metrics: simple, easy and tempting

ABSTRACT

The launch of Google Scholar Citations and Google Scholar Metrics may provoke a revolution in the research evaluation field as it places within every researcher's reach tools that allow bibliometric measuring. In order to alert the research community over how easily one can manipulate the data and bibliometric indicators offered by Google's products we present an experiment in which we manipulate the Google Citations' profiles of a research group through the creation of false documents that cite their documents, and consequently, the journals in which they have published modifying their H-index. For this purpose we created six documents authored by a faked author and we uploaded them to a researcher's personal website under the University of Granada's domain. The result of the experiment meant an increase of 774 citations in 129 papers (six citations per paper) increasing the authors and journals' H-index. We analyse the malicious effect this type of practices can cause to Google Scholar Citations and Google Scholar Metrics. Finally, we conclude with several deliberations over the effects these malpractices may have and the lack of control tools these tools offer.

KEYWORDS

Google Citations / Google Scholar Metrics/ Scientific Journals / Scientific fraud / Citation analysis / Bibliometrics / H Index / Evaluation / Researchers

Referencia bibliográfica recomendada

Delgado López-Cózar, Emilio; Robinson-García, Nicolás; Torres Salinas, Daniel (2012). Manipular Google Scholar Citations y Google Scholar Metrics: Simple, sencillo y tentador. *EC3 Working Papers* 6: 29 de mayo de 2012

1. INTRODUCCIÓN

Si el nacimiento de Google Scholar en 2004, - un novedoso buscador para recuperar todo tipo de documentos académicos y de las citas asociadas a los mismos, - revolucionó el mercado de la información científica permitiendo el acceso universal y gratuito a todos los documentos circulantes en la web, el lanzamiento de Google Scholar Citations, una herramienta de medición de la producción y del impacto de los investigadores (Cabezas-Clavijo y Torres-Salinas, 2012), y de Google Scholar Metrics, un índice de impacto científico de revistas a partir del recuento de citas (Cabezas-Clavijo y Delgado López-Cózar, 2012), pueden suponer un innegable hito en la globalización y democratización de las evaluaciones bibliométricas (Butler 2011). Aparte de constituir un auténtico órdago a las tradicionales bases de datos bibliográficas e índices bibliométricos patrocinados por Thomson Reuters (Web of Science y JCR) y Elsevier (Scopus y SJR) y de romper el duopolio de facto ejercido por estas compañías convirtiéndose en el más inmediato futuro en un serio competidor, los nuevos productos de Google Scholar proyectan unos escenarios y unos dilemas de orden ético y sociológico de dimensiones colosales e imprevistas consecuencias para el mundo de la ciencia y de su evaluación.

Independientemente de los problemas técnicos y metodológicos de que adolece la familia Google Scholar, sobre los que se está investigando y discutiendo (Jacsó, 2008, 2011; Wouters y Costas, 2012; Aguillo, 2012; Cabezas-Clavijo y Delgado López-Cózar, 2012; Torres-Salinas, Ruiz-Pérez y Delgado López-Cózar, 2009) y que seguramente puedan ser resueltos, su irrupción supone una auténtica voladura de todos los controles o filtros científicos a los que estaban sometidos los investigadores, sus trabajos y las publicaciones en las que comunicaban sus resultados de investigación, constituyendo un nuevo desafío para el mundo de la evaluación. Desde el momento en que Google Scholar automáticamente rastrea, indiza y vacía cualquier documento de apariencia científica que cuelga de un dominio académico por voluntad de un autor sin sufrir ningún control externo previo (el de los repositorios es solo un filtro técnico que no opera sobre los contenidos) abre la posibilidad a que cualquier persona sin escrúpulos pueda manipular a su entero gusto y beneficio documentos que repercuten directamente sobre la evaluación de su producción e impacto bibliométrico.

Y puesto que modificar o manipular la producción o el impacto de los trabajos mediante la auto-citación controlada intencionadamente o de forma desmedida es algo plenamente vigente, consideramos más que oportuna la necesidad de analizar en profundidad la capacidad de Google para detectar la manipulación de datos.

Este trabajo sigue la línea abierta por Labbé (2010) con aquel experimento mediante el que logró convertir al inexistente investigador Ike Antkare (cuya pronunciación en inglés equivaldría a la frase '*I can't care*', me da igual) en uno de los científicos más citados de la historia. No obstante, en este caso queremos incidir en aquellos aspectos que pueden resultar más peligrosos de cara al uso de estas herramientas con fines evaluadores y de los efectos perversos que pueden tener en el comportamiento de los investigadores. Por tanto, el objetivo que se persigue es demostrar lo fácil que es manipular los indicadores bibliométricos de Google Scholar y de sus productos derivados. Pero, a diferencia de Labbé, queremos incidir no sólo en los aspectos más técnicos, sino también en la dimensión más sociológica y de la enorme tentación que puede suponer para investigadores y editores de revistas caer en el falseo de citas para incrementar su impacto. Para ello, mostraremos cómo se pueden manipular los indicadores bibliométricos de varios autores y revistas simultáneamente y de la forma más sencilla posible: colgando en nuestra misma página web varios documentos ficticios que se limitan a reproducir toda la producción bibliográfica de todo un grupo de investigación. No es necesario recurrir a un software específico para la creación de documentos falsos: basta con el corta y pega y la generación de un documento HTML colgado en nuestra página web bajo dominio institucional. Asimismo, pretendemos evidenciar cual es la capacidad de Google para detectar la manipulación, y ponerle remedio, así como para eliminar tanto los documentos manipulados como las citas que emiten.

Este trabajo experimental de fabricación de documentos ficticios o falsos para poner en evidencia los defectos, sesgos o errores cometidos por los investigadores y su empresas científicas es un recurso habitual, especialmente en el ámbito de la evaluación de la ciencia. Baste mencionar las célebres falsificaciones de Peters & Ceci (1990), Epstein (1990), Sokal (1996, 1997) o Baxt et al. (1998) para demostrar las insuficiencias del peer review como método objetivo, fiable, válido, eficiente y libre de sesgos de control de la calidad de los contenidos publicados en las revistas científicas; o el más reciente Scigen¹, un programa elaborado por tres estudiantes del MIT para generar trabajos aleatorios en el campo de la Informática incluyendo gráficos, figuras y citas; trabajos todos ellos que levantaron ampollas y desataron agrias polémicas dentro de la comunidad científica.

Así pues, este trabajo está estructurado de la siguiente manera. En un primer lugar describimos la metodología que se siguió para realizar el experimento, cómo se crearon y dónde se alojaron los documentos falsos. A continuación, mostramos el efecto que produjo en los perfiles bibliométricos de los investigadores afectados por las citas emitidas por los documentos falsos y emulamos la incidencia que tendría en las revistas afectadas si Google Scholar Metrics se actualizara del mismo modo que el resto de productos. Analizamos las incidencias de tipo técnico detectadas y los peligros que plantean estas herramienta para la evaluación. Finalmente concluimos resaltando los puntos fuertes de las mismas y con una serie de consideraciones finales.

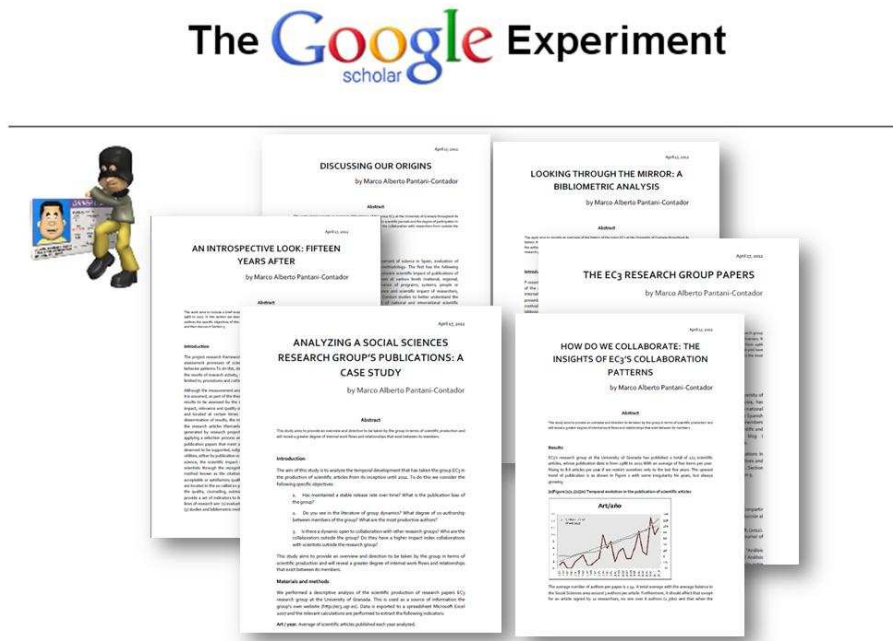
2. LA MANIPULACIÓN DE DATOS: EL EXPERIMENTO GOOGLE SCHOLAR

Con el objetivo de analizar la capacidad de Google Citations de discernir los trabajos de investigación de los que no lo son y de comprobar la dificultad que pudiera entrañar

¹ <http://pdos.csail.mit.edu/scigen/>

manipular la producción y la citación de los documentos registrados en Google Scholar y en sus productos bibliométricos (GSC, GSM), decidimos crear unos documentos artificiales que reprodujeran las referencias bibliográficas a todos los trabajos de todos los miembros del grupo EC3 (Evaluación de la Ciencia y de la Comunicación Científica), disponibles en <http://ec3.ugr.es> de la manera más arcaica posible, apta para los menos doctos en cuestiones técnicas. De esta forma, pretendemos demostrar cómo está al alcance de cualquier investigador que se lo proponga falsear la producción y la citación recogidas por Google Citations.

Figura 1. Documentos falsos adjudicados a MA Pantani-Contador



Siguiendo el ejemplo del trabajo de Labbé (2010), creamos un investigador falso al que decidimos llamar Marco Alberto Pantani-Contador, haciendo alusión tanto al gran fraude de dopaje del ciclista italiano, como a la circunstancia supuestamente accidental que privó al español de su Tour. Así, Pantani-Contador produjo seis documentos (figura 1) que no pretendían ser tomados como publicaciones científicas en ningún momento, sino que se explicitó que se trataban de *working papers*. En una operación que duró no más de una tarde de trabajo, se redactó mal y pronto un pequeño texto en que se incluían gráficos y tablas así como texto copiado y pegado de la página web del grupo para posteriormente traducirlo todo de manera automática con Google Translate y trocearlo en seis documentos distintos. Cada documento emite 129 citas que corresponden a 129 trabajos producidos por los distintos miembros del grupo EC3 de acuerdo con los datos de la web <http://ec3.ugr.es>. Es decir, que el incremento de citas esperado es de 774 citas.

Hecho esto, se elaboró una sencilla página web bajo el dominio de la Universidad de Granada en la que se incluyeron los mismos y se enlazaron al texto completo, con la intención de que fueran indexados por Google Scholar. Se excluyeron otros servicios para el alojamiento de documentos académicos como son los repositorios institucionales y temáticos ya que éstos no están obligados a hacer ningún tipo de control bibliográfico

más allá del formal (Delgado López-Cózar, 2012) y además se escapaban del objeto de este experimento, que era la manipulación directa de citas.

Los documentos falseados se subieron a la web el 17 de abril de 2012. Presumiblemente debido a que se trataba de una página web personal y no de un repositorio, la indexación por parte de Google de los documentos falsos no se llevó a cabo hasta el 12 de mayo de 2012, es decir, aproximadamente un mes después. En ese momento, los distintos miembros del grupo así como los tres coautores de los trabajos publicados recibieron la alerta correspondiente de GS Citations indicándoles que un tal MA Pantani-Contador había citado sus trabajos. La explosión de citas fue apabullante, en especial entre los investigadores con una menor trayectoria investigadora donde el número de citas de multiplicó por seis, engrosando notoriamente su perfil investigador en Google Citations.

Figura 2. Incremento de citas de los autores de este trabajo

		Emilio Delgado López-Cózar		
		TODO EL PERIODO		
		ANTES del experimento	DESPUÉS del experimento	
Citations		862	1297	+ 435
H-Index		15	17	+ 2
i10-Index		20	40	+ 20

		DESDE 2007		
		ANTES del experimento	DESPUÉS del experimento	
Citations		560	995	+ 435
H-Index		10	15	+ 5
i10-Index		11	33	+ 22

		Nicolás Robinson-García		
		TODO EL PERIODO		
		ANTES del experimento	DESPUÉS del experimento	
Citations		4	29	+ 25
H-Index		1	4	+ 3
i10-Index		0	0	0

		DESDE 2007		
		ANTES del experimento	DESPUÉS del experimento	
Citations		4	29	+ 25
H-Index		1	4	+ 3
i10-Index		0	0	0

		Daniel Torres-Salinas		
		TODO EL PERIODO		
		ANTES del experimento	DESPUÉS del experimento	
Citations		227	409	+ 182
H-Index		9	11	+ 2
i10-Index		7	17	+ 10

		DESDE 2007		
		ANTES del experimento	DESPUÉS del experimento	
Citations		226	408	+ 182
H-Index		9	11	+ 2
i10-Index		7	17	+ 10

En la figura 2 observamos cómo prácticamente el incremento de citas que sufren los perfiles coincide con el número de citas inyectado. Obviamente, la incidencia dependerá del número de publicaciones que tenga cada autor del grupo de investigación objeto de estudio. Así, el mayor incremento de citas lo recibe quien menos tenía de partida, en este caso, Robinson-García pasa a multiplicar por 7,25 el número de citas, mientras que Torres-Salinas casi las dobla y las citas recibidas de Delgado López-Cózar se incrementa en 1,5. También se observa el efecto que tiene en el índice H de cada investigador. Mientras que el incremento es notable en el perfil del investigador con menor producción científica, la variación para los otros dos investigadores resulta mucho más moderada,

ilustrando la estabilidad del indicador. Nótese el caso de Torres-Salinas donde se dobla el número de citas mientras que el índice H sólo se ve incrementado en dos. Por otra parte, observamos cómo el i10-index se ve afectado notablemente para los investigadores más experimentados. En el caso de Torres-Salinas se incrementa más del doble pasando de 7 a 17 y en el caso de Delgado López-Cózar se incrementa el triple para los últimos cinco años, pasando de 11 a 33.

Figura 3. Vista de la manipulación en el perfil de uno de los autores



Por otra parte, resulta interesante analizar el efecto que pudiera tener el incremento de citas en el índice H de las revistas indexadas en el Google Scholar Metrics. Para ello, hemos considerado las dos revistas en las que los miembros del grupo de investigación objeto de estudio ha publicado más trabajos y por tanto, más susceptibles de manipulación. Estas son El Profesional de la Información con 30 trabajos publicados en esta revista y la Revista Española de Documentación Científica, con 33. En la tabla 1 mostramos el índice H que tienen El Profesional de la Información y la Revista Española de Documentación Científica de acuerdo con los datos publicados por Google y el incremento que se produciría al añadir las citas emitidas por el ficticio Pantani-Contador. Hemos de recordar que este producto de Google, al contrario que el resto, no se actualiza automáticamente y que los datos relativos al mismo corresponden a los de su lanzamiento, es decir, al 1 de abril de 2012 (Cabezas-Clavijo y Delgado López-Cózar, 2012). Observamos cómo la mayor variación la sufriría El Profesional de la Información, donde siete trabajos superarían el umbral de las 12 citas incrementando notablemente su índice

H y ascendería por tanto en el ranking correspondiente a las revistas escritas en español de la posición 20 en la que se encuentra actualmente, a la posición 5 si el índice se actualizara y publicara hoy mismo. Por su parte, la Revista Española de Documentación Científica sufriría una alteración mínima, ya que solo un artículo supera el umbral de las 9 citas que suponen una alteración en su índice H. Aún así y debido al gran número de revistas que tienen su mismo índice H, ascendería veinte puestos de la posición 74 a la posición 54.

Tabla 1. Efecto del falseo de citas sobre las revistas

Revista	Índice H (GS Metrics)	Art > 12 citas	Índice H manipulado
El Profesional de la Información	12	7	19
Revista Española de Documentación Científica	9	1	10

Tras comprobar la fragilidad de los productos de Google Scholar de cara a la inclusión de los documentos falsos y mostrar los efectos nocivos que tiene tanto a nivel investigador como a nivel de revista, el 17 de mayo retiramos tanto los documentos como la web donde estaban reseñados con el objeto de ver si eran posteriormente retirados de Google Scholar y consecuentemente, las citas falsas emitidas a los investigadores objeto de estudio se eliminaban volviendo todo a su cauce. A día de hoy (29 de mayo de 2012) 17 días después de haberse retirado los documentos falsos de la web, siguen apareciendo tanto los documentos falsos en Google Scholar como las citas falsas en los perfiles de GS Citations de los investigadores. Los registros siguen apareciendo al realizar una búsqueda por el autor ficticio y, aunque los enlaces están rotos, está disponible la versión guardada en caché.

3. CONSIDERACIONES DE TIPO TÉCNICO

Ante los resultados expuestos, se hace más que obvia la capacidad de alterar de manera rápida y sencilla los perfiles de citación que ofrece Google. Esto pone de relieve el peligro que puede entrañar poner al alcance de la mano de editores e investigadores la tentación de hacer ‘ingeniería de citas’ y dopar su índice H engordando artificialmente las citas de todos los documentos o, siendo mucho más sofisticado, dirigiendo las citas exclusivamente a la zona caliente de las publicaciones, es decir, a aquellas que están en la frontera para influir el índice H de un autor o una revista. En el caso de la revista El Profesional de la Información, son 16 los documentos localizados con entre 10 y 12 citas para el periodo analizado por Google Scholar Metrics (2007-2011), esto supone que añadiéndole entre 1 y 3 citas a cada uno de estos documentos, la revista podría ascender hasta situarse en las primeras posiciones del ranking ofrecido para las revistas en español.

Volviendo a cuestiones más técnicas, en primer lugar, destacar la facilidad para manipular la producción, algo denunciado en su momento por Labbé (2010), pero también las citas de los autores. Esto resulta cuanto menos preocupante sobre todo por la poca capacidad de Google Scholar para discernir los documentos falsos de los verdaderos, tanto en cuanto que se trata de una herramienta inicialmente creada para indexar y recuperar todo tipo de documentos de carácter científico en su sentido más amplio. La inclusión de herramientas de evaluación de la actividad investigadora como son GS Citations y GS Metrics debe conllevar consigo la introducción de herramientas de control y establecer criterios menos

laxos en la indexación de los documentos. En este sentido, Google Scholar ofrece acceso a una amplia tipología documental que lo hace mucho más atractivo frente a las bases de datos convencionales no solo por su 'fórmula mágica' para la recuperación de información, sino también por la riqueza de la información que maneja. Sin embargo, abandonar un entorno tan controlado como el de las revistas, entraña sus peligros en el mundo de la evaluación.

Por otra parte, resulta más que interesante comprobar la estabilidad del índice H sobre todo cuando afecta a autores con una carrera investigadora de cierta trascendencia, aún doblando en algún caso el número de citas, la variación para este indicador era mínima, lo que puede trasladar cierta tranquilidad, sin embargo, lamentablemente son muchas y variadas las formas de manipulación de la auto-citación que pueden incidir en este indicador (Bartneck y Kokkelmans, 2011) llamando a la cautela. Además, en lo que respecta a las revistas y con las más que presumibles actualizaciones de GS Metrics, que hace pocos días pasó a estar disponible en la portada de la versión inglesa de Google Scholar, puede hacer relativamente fácil para editores maliciosos la modificación de sus índices H. Por otra parte, se observa cómo la variación del i10-Index es notable en este caso, cuanto mayor sea la trayectoria investigadora del investigador.

En lo relativo al efecto que pueden tener estas prácticas en los rankings de revistas presentados por los de Mountain View, obviamente sería significativo, al menos en el caso de las revistas donde aún abundan los pequeños números y la más ínfima variación puede disparar a lo más alto o hundir en la miseria a la revista en cuestión.

Figura 4. Pantalla de resultados de Google Scholar

The screenshot shows the Google Scholar search interface. At the top, the Google logo is on the left, and a search bar contains the text 'author:"MA Pantani"'. To the right of the search bar is a magnifying glass icon. Below the search bar, the word 'Scholar' is displayed in red, followed by '6 results (0.07 sec)'. On the far right, there is a 'My Citations' button with a pencil icon. The search results are listed under the heading 'Articles'. A red arrow points to the first result, which is a PDF document titled 'HOW DO WE COLLAB INSIGHTS OF EC3'S CO' by MA Pantani, dated 2012. The abstract for this document reads: 'EC3's research group at the University articles, whose publication date is from 1 Rising to 8.6 articles per year if we restri trend of publication is as shown in Figu growing. ... The average number of authors per pape the Social Sciences area around 3 authors for an article ...'. Below this, there are three more results, each with a PDF icon and a title: 'ANALYZING A SOCIAL SCIENCES RESEARCH GROUP'S PUBLICATIONS: A CASE STUDY', 'THE EC3 RESEARCH GROUP PAPERS', and 'LOOKING THROUGH THE MIRROR: A BIBLIOMETRIC ANALYSIS'. Each result also includes the author's name (MA Pantani-Contador) and the year (2012). The abstract for the last result reads: 'Abstract This work aims to provide an overview of the history of the group EC3 at the University of Granada throughout its history. It looks at developments in the publication of articles in scientific journals and the degree of participation in the authorshio of the work ...'. On the left side of the results, there are filters for 'Legal documents', 'Any time' (with options for 'Since 2012', 'Since 2011', 'Since 2008', and 'Custom range...'), and checkboxes for 'include patents' and 'include citations', both of which are checked.

La imposibilidad de editar las citas recibidas en GS Citations señalando las erróneas e indicándole al sistema las ausentes, remarca aún más esta importante limitación del

sistema, alertando como ya se hizo en su momento, de los peligros que entraña un uso descuidado de esta herramienta con fines bibliométricos (Cabezas-Clavijo y Torres-Salinas, 2012). Además, queda por ver si llegan a eliminarse los registros de los trabajos de nuestro inexistente investigador y con ellos, las citas falsas emitidas. Hasta la fecha no ha sido así, y de no producirse, dejará patente una realidad ya presente en el buscador general, la imposibilidad de ejercer el 'derecho al olvido' (Gómez, 2011).

Ahora bien, conviene recalcar que la visibilidad que ofrecen estas herramientas y la transparencia que permiten hacen que rápidamente este tipo de prácticas sean fácilmente detectadas por la comunidad, tal y como hemos presenciado durante la realización de este experimento con la llamada de colegas y compañeros preguntando y comentando el suceso tras recibir el pertinente correo de alerta de GS Citations.

Por otra parte, resulta interesante observar cómo documentos iguales y creados sobre una misma plantilla son indexados de distinta manera por Google, Poniendo nuevamente de relevancia los escasos procesos de normalización que realiza cuando procesa los datos de los documentos que se suben a la web. Así, observamos cómo de los seis documentos falsos que se crearon y fueron posteriormente indizados, Google olvidó el segundo apellido de nuestro investigador 'fantasma' (figura4).

3. REFLEXIONES FINALES Y CONCLUSIONES

Si en trabajos anteriores hemos apostado abiertamente por la utilización de Google Scholar para la evaluación minimizando los efectos de los sesgos y problemas técnicos y metodológicos (Cabezas-Clavijo, Delgado López-Cózar, 2012), con este trabajo queremos alertar a la comunidad científica de la facilidad que existe para manipular los datos e indicadores bibliométricos. Pasar de un entorno donde la producción, difusión y evaluación del conocimiento científico estaba controlada (con todas las limitaciones que se le quieran atribuir) a otro en el que no existe ningún control más que el que dicta la propia conciencia del investigador es un novedad radical que encierra numerosos peligros (Tabla 2).

Tabla 2. Medidas de control en el modelo tradicional vs. productos Google Scholar

Modelo tradicional	Productos Google Scholar
Las bases de datos seleccionan las revistas a indexar	Indexa todo documento proveniente de un dominio académico o científico
Las revistas seleccionan los trabajos a publicar	Cualquier documento indexado emite y recibe citas.
Existe un control de enlazado entre citante y citado Se persiguen los comportamientos fraudulentos	No es posible denunciar comportamientos fraudulentos o errores en el enlazado de citas

Poner en manos de todos los científicos, que son humanos, las herramientas que permiten manipular la producción y las citas pueden tener consecuencias imprevisibles o pueden convertir a estas herramientas en inservibles. La falta de control de estos productos son su punto fuerte y su talón de Aquiles. La excesiva facilidad que permite en su manejo GS Citations hace por contra, que cualquiera que se lo proponga pueda crearse su propio perfil bibliométrico ficticio y, emulando a Ike Antkare, convertirse en el científico más productivo y con mayor impacto de su especialidad. Cuanto menos los editores de revistas si finalmente Google Scholar Metrics se incorpora definitivamente, que también pueden verse tentados a incrementar el impacto de sus revistas utilizando técnicas de dudosa ética.

Estos productos accesibles y gratuitos, aparte de despertar el narciso que habita en los científicos (Wouters; Costas, 2012), pueden exacerbar las malas artes y prácticas comunicativas de los científicos a fin de manipular la orientación y el sentido de los números como resultado de la creciente presión por publicar avivada por los sistemas de evaluación del rendimiento científico adoptados en todos ámbitos y países. Los casos de fraudes de editores que se aprestan a manipular las políticas editoriales de sus revistas con el fin de incrementar el factor de impacto han sido descritas minuciosamente por Falagas y Alexiou (2008). Son abundantes los casos y de hecho, todos los años anuncia Web of Science las revistas expulsadas por malas prácticas de sus bases de datos (http://admin-apps.webofknowledge.com/JCR/static_html/notices/notices.htm). Ejemplos no faltan, como el expuesto en el trabajo de Dimitrov et al. (2010) con el sonado caso de la revista *Acta Crystallographica A* que sorprendió a propios y extraños pasando su factor de impacto de 2,38 a 49,93 de un año para otro. Sin embargo, resultó que de las 5966 citas que recibieron en 2009 sus 72 artículos publicados en 2008, solo uno de ellos acaparaba 5624 citas. Este artículo fue precisamente el responsable de semejante fenómeno. a la búsqueda de incrementar la repercusión Otro ejemplo lo encontramos en el trabajo de Opatrný (2008).

A día de hoy no existen más controles o filtros para evitar el fraude en los datos que las reservas éticas y morales de los propios investigadores. En este sentido, cabe destacar la labor de entidades como el Committee on Publication Ethics (<http://publicationethics.org/>) y otras entidades similares que persiguen el fraude y el falseo en el mundo científico tradicional, el de la revista. Tal vez estamos en ciernes de una nueva revolución en el modelo de comunicación científica y como tal, solo es cuestión de tiempo la aparición de estamentos similares a estos. Nosotros damos por concluida nuestra experimentación y esperamos con paciencia a que se retire del mercado a nuestro investigador fantasma esperando que, siguiendo nuestro ejemplo, Google elimine las citas falsas con la esperanza de que ese mínimo control exista. el esfuerzo de Google en la producción de estas herramientas vaticina un cambio de paradigma en el mundo de la evaluación de la investigación. No solo por su gratuidad, sino también por su gran cobertura, inmediatez y facilidad de uso. Pero queda por ver cuál será el camino que tome Google para poner diques a esos números que devoran la ciencia (Monastersky 2005).

MATERIAL COMPLEMENTARIO

Para más información sobre los resultados del experimento, consulte la web: <http://www.ugr.es/~elrobin/pantani.html>.

BIBLIOGRAFÍA

Aguillo, I. (2012). Is Google Scholar useful for bibliometrics? A webometric analysis. *Scientometrics* 91; 2: 343-351.

Bartneck, C.; Kokkelmans, S. (2011). Detecting h-index manipulation through self-citation analysis. *Scientometrics* 87; 1: 85-98.

Baxt, W. G.; Waeckerle, J. F.; Berlin, J. A.; Callahan, M. L. (1998). Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer review performance. *Annals of Emergency Medicine* 32; 3: 310-317.

Butler, D. (2011). Computing giants launch free science metrics. *Nature* 476; 18: doi:[10.1038/476018a 2](https://doi.org/10.1038/476018a2)

Cabezas-Clavijo, Á; Delgado López-Cózar, E. (2012). El impacto de las revistas según Google, ¿un divertimento o un producto científico aceptable? *EC3 Working Papers* 1. accesible en <http://eprints.rclis.org/handle/10760/16836>

Delgado López-Cózar, E. (2012). Los repositorios en Google Scholar Metrics o ¿qué hace un tipo documental como tú en un lugar como ese? *EC3 Working Papers* 4: 3 de abril de 2012.

Dimitriv, J. D.; Kaberi, S. R.; Bayry, J. (2010). Metrics: journal's impact factor skewed by a single paper. *Nature* 466: 179,

Epstein, W. M. (1990). Confirmational response bias among social work journals. *Science, Technology & Human Values* 15; 1: 9–38.

Falagas, M. E. y Alexiou, V. G. (2008). The top-ten in journal impact factor manipulation. *Archivum Immunologiae et Therapiae Experimentalis* 56; 4: 223-226

Gómez, R. G. (2011). Quiero que Internet se olvide de mí. *El País* 7 de enero, 2011.

Jacsó, P. (2008). The pros and cons of computing the h-index using Google Scholar. *Online Information Review* 32; 3: 437-452.

Jacsó, P. (2011). Google Scholar duped and deduped – the aura of ”robometrics”. *Online Information Review* 35; 1: 154-160.

Labbé, C. (2010). Ike Antkare, one of the greatest stars in the scientific firmament. *ISSI Newsletter* 6; 1: 48-52.

Monastersky R. (2005)The Number That’s Devouring Science. *The Chronicle of Higher Education* 52; 8: A12. Disponible en: <http://chronicle.com/free/v52/i08/08a01201.htm>

Opatrný, T. (2008). Playing the system to give low-impact journal more clout. *Nature* 455: 167.

Peters, D. P.; Ceci, S. J. (1990). Peer-review practices of psychological journals – the fate of accepted, published articles, submitted again. *Behavioral and Brain Sciences*, 5; 2: 187-195.

Sokal, A. D.; Bricmont, J. (1997). *Impostures Intellectuelles*. Editions Odile Jacob.

Wouters, P.; Costas, R. (2012). *Users, narcissism and control – tracking the impact of scholarly publications in the 21st century*. SURFfoundation. Accesible en: <http://www.surf.nl/nl/publicaties/Documents/Users%20narcissism%20and%20control.pdf>