

¿Es reproducible el *Altmetric Attention Score*?

Is the *Altmetric Attention Score* reproducible?

José-Luis Ortega

Ortega, José-Luis (2019). "¿Es reproducible el *Altmetric Attention Score*?". *Anuario ThinkEPI*, v. 13, e13e08.

<https://doi.org/10.3145/thinkepi.2019.e13e08>

Publicado en *IweTel* el 28 de noviembre de 2019

José-Luis Ortega

<https://orcid.org/0000-0001-9857-1511>

Institute for Advanced Social Studies (IESA-CSIC)
Joint Research Unit Knowledge Transfer and
Innovation (UCO-CSIC)
Córdoba, Spain
jortega@iesa.csic.es



Resumen: La presente nota responde de una forma empírica a la cuestión de en qué medida el *Altmetric Attention Score* (AAScore) puede ser reproducido. Para ello se tomó una muestra de más de 28.000 artículos extraídos de la API pública de *Altmetric.com* para reproducir el indicador a partir de los pesos publicados por la plataforma y se compararon con los valores reales. Sólo en el 13,2% de los casos se replicó correctamente el AAScore, incrementándose los errores a medida que aumenta el impacto. Las métricas que más influyen en su no reproducibilidad son *Tweeters* (44%), news (16%) y blogs (12%). La principal razón de esta pobre reproducibilidad es la existencia de criterios de puntuación no publicados que impiden su replicación. Este hecho limita considerablemente el uso del AAScore en procesos de evaluación científica.

Palabras clave: Bibliometría; Altmetría; *Altmetric Attention Score*; indicadores científicos; Reproducibilidad.

Abstract: The aim of this study is to know to what extent the *Altmetric Attention Score* (AAScore) can be reproduced. A sample of more than 28,000 research articles extracted from the public API of *Altmetric.com* were employed to reproduce the AAScore from the weights published in *Altmetric.com*. The indicator was correctly replicated in 13.2% of the cases, increasing the errors as the impact increases. The metrics that influence the most on the non-reproducibility are *Tweeters* (44%), news (16%) and blogs (12%). The main reason of this problem is the existence of non-published scoring criteria that impede the replication. This fact is an important limitation to the use of AAScore as research evaluation indicator.

Keywords: Bibliometrics; Altmetrics; *Altmetric Attention Score*; research indicators; Reproducibility.

1. Introducción

Altmetric.com es uno de los proveedores de métricas alternativas más populares y el más utilizado en la comunidad académica (Ortega, 2018). Este éxito se debe, entre otras cosas, a la política abierta de esta empresa con respecto al uso científico de sus datos (*Altmetric Annual Research Grant*, *Researcher Data Access Program*, API pública). Uno de los elementos que más ha servido para el éxito de esta compañía ha sido el diseño de un indicador que reúne en un único guarismo todo el impacto mediático que recibe la producción científica. De esta forma, es fácil identificar los trabajos que más huella han dejado en la Web (*Altmetric Top 100*) y qué posición ocupan estos dentro del conjunto de trabajos recopilados por la plataforma (*Attention Score in context*).

El *Altmetric Attention Score* (AAScore) es un indicador lineal que agrega las diferentes métricas que la plataforma captura referente a un trabajo (exceptuando lectores de *Mendeley*, guardados de *CiteULike* y citas bibliográficas). A cada métrica se le asigna un peso definido según la importancia de cada una. De esta forma, cuantas más menciones tenga una publicación, mayor será su puntuación. Sin embargo, este indicador ha sido objeto de duras críticas:

- agrupa métricas sin tener en cuenta su significado, mezclando diferentes tipos de impactos;
- los pesos son arbitrarios y no están basados en criterios científicos (**Gumpenberger; Glänzel; Gorraiz, 2016; Mukherjee; Subotić; Chaubey, 2018**);
- los valores son redondeados (en el *Altmetric Explorer*), causando imprecisiones en las puntuaciones más bajas.

Sin embargo, una de las cuestiones más importantes para la validez de un indicador como instrumento de evaluación científica es que permita su reproducción por aquellos que son evaluados (**Hicks et al., 2015**). En la presente nota se pretende hacer un pequeño estudio que nos permita conocer en qué grado este indicador puede ser reproducido a partir de la información que *Altmetric.com* proporciona para su cómputo (*NISO, 2016*).

Se seleccionó una muestra de 28.123 artículos extraídos de *Altmetric.com* en mayo de 2017. <https://api.altmetric.com>

Se trata de una muestra proveniente de *PlumX* y limitada a fecha de publicación posterior a 2013 (**Ortega, 2018**). A continuación, se calculó el AAScore a partir de la lista de métricas y sus correspondientes pesos publicada por *Altmetric* para el cálculo de su indicador (*Altmetric Support, 2019*). Puesto que los datos fueron tomados en 2018, la métrica *Patent* y su peso no fue considerada para el cómputo, ya que en 2018 esta métrica aún no estaba incluida en *Altmetric.com*. Para una mejor comparación entre el AAScore y el calculado, se ha tomado el AAScore sin redondear que proporciona la API pública de *Altmetric.com*.

2. Resultados

Los resultados muestran que solo en un 13,2% de los casos se ha podido reproducir el indicador con el mismo resultado. Esto es, el 86,8% del AAScore calculado no coincide con el AAScore publicado. Las mayores coincidencias se dan en las puntuaciones 1 (44,1%) y 0,25 (29,7%), lo que demuestra que las coincidencias son más frecuentes en valores bajos que en altos.

La figura 1 confirma esta percepción al mostrar la relación entre los valores que AAScore proporciona y las diferencias o errores de estos con respecto a los calculados. Como se puede observar, hay una relación lineal entre los valores reales y los errores, lo que significa que cuanto mayor es el valor que adquiere el AAScore, mayor es el error a la hora de reproducir el indicador. Es también importante señalar que la mayoría de los errores toma valores positivos, lo que muestra que los valores calculados son más elevados que los valores reales de *Altmetric.com*.

La razón de esta enorme discrepancia tanto en el número de errores como en su magnitud, está en que muchos de estos pesos son alterados por *Altmetric.com* siguiendo criterios particulares que impiden conocer el peso real de cada indicador. Por ejemplo, las menciones de noticias (*news*) son valoradas de

“Los resultados muestran que solo en un 13,2% de los casos se ha podido reproducir el indicador con el mismo resultado”

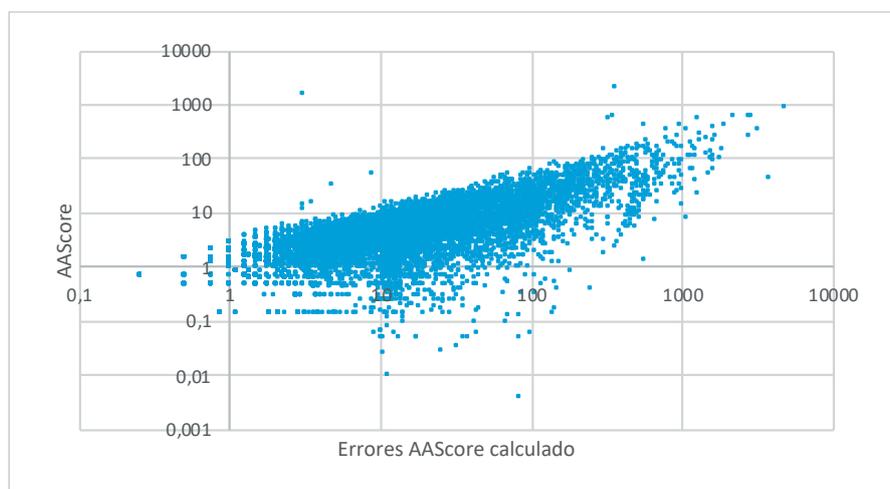


Figura 1. Relación entre el AAScore y las diferencias con el AAScore calculado (log-log)

forma diferente según el alcance del medio, de forma que

“una mención desde un popular medio de ámbito nacional como *The New York Times* aportará más a la puntuación que una mención desde un medio pequeño y especializado como *2Minute Medicine*” (traducción del autor).

Esta práctica se hace extensible a otros indicadores como *Tweeters*, *Policy documents*, *Patent Citations*, *Open Syllabus* and *Wikipedia*, donde el peso es modificado por criterios desconocidos por el público. En algunos indicadores como *Wikipedia* y *Open Syllabus* el peso siempre es fijo, no se incrementa por ser mencionado varias veces; en el caso de *Tweeters*, *Policy documents* o *Patent Citations* los pesos son alterados según distinciones dentro de la propia métrica que *Altmetric.com* conoce, pero no publica. Por ejemplo, los retweets tienen un peso diferente (0,85), pero no se puede calcular ya que la plataforma no distingue los tweets originales de los retweets. Además, en esta misma métrica, las puntuaciones son modificadas de acuerdo a tres elementos: alcance, promiscuidad, sesgo. Estos elementos no son cuantificados y por lo tanto no sabemos cómo influye en el valor final de un tweet.

La figura 2 muestra los porcentajes de *AAScore* reproducidos satisfactoriamente y los no distribuidos por las métricas implicadas en su cálculo. Así, por ejemplo, el 97,2% de los *AAScore* no reproducidos tenían algún tweet, mientras que en los sí reproducidos el porcentaje desciende al 52,9%. Esto quiere decir que cuando un artículo ha sido tuiteado, la probabilidad de que su *AAScore* pueda ser reproducido desciende a algo menos de la mitad (44,3%). En el caso de blogs y news esta diferencia se produce en un 16% en el caso de news y un 11,7% en el de blogs. En el resto de métricas, las diferencias no son significativas.

La respuesta más clara a la pregunta inicial es que *AAScore* solo es un indicador reproducible en el 13% de los casos. Este problema afecta más a las puntuaciones altas que a las bajas, las cuales tienen menos parámetros para su cálculo y su reproducibilidad es por tanto más probable. El error de cálculo aumenta según se incrementa la puntuación, haciendo casi imposible obtener un valor aproximado en artículos con un destacado impacto. Por tipo de métricas, *Tweeters*, blogs y news son las que más influyen en que no se pueda reproducir el indicador. Las razones de su baja reproducibilidad se deben fundamentalmente a que los valores de los pesos de cada métrica están condicionados a criterios y valoraciones desconocidos para los usuarios. Se puede concluir, entonces, que dado el bajo porcentaje de reproducibilidad el *AAScore* no puede ser aceptado como indicador válido para la evaluación científica, ya que una de las condiciones más importantes en los procesos de evaluación es que los investigadores puedan conocer cómo se han calculado los indicadores usados en su evaluación (*DORA*, 2013; *Hicks et al.*, 2015). Una forma de mejorar su reproducibilidad podría ser la publicación del ranking de valoración de medios y blogs, y los criterios de puntuación de tweets y retweets.

“Cuanto mayor es el valor que adquiere el *AAScore*, mayor es el error a la hora de reproducir el indicador”

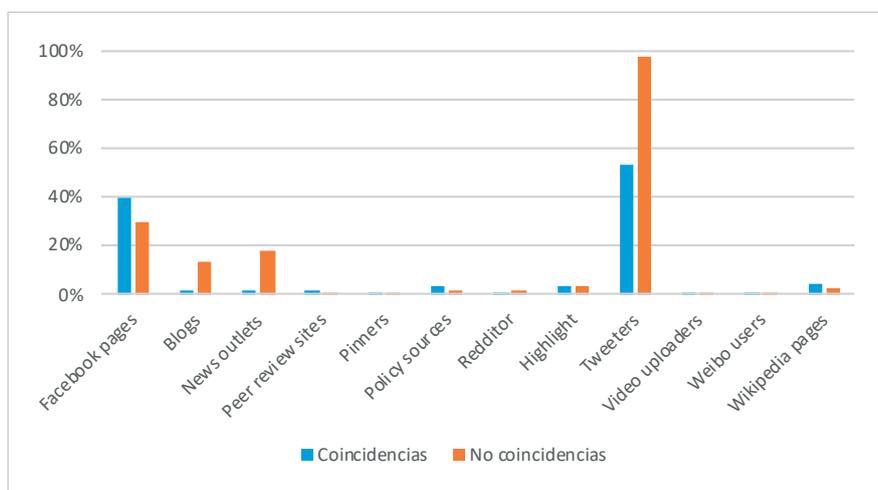


Figura 2. Porcentaje de *AAScore* reproducidos y no reproducidos por métricas

“La razón de esta enorme discrepancia tanto en el número de errores como en su magnitud, está en que muchos de estos pesos son alterados por *Altmetric.com* siguiendo criterios particulares que impiden conocer el peso real de cada indicador”

3. Referencias

Altmetric Support (2019). "How is the Altmetric Attention Score calculated?". *Altmetric support*, 29 Jul. <https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-score-calculated->

American Society for Cell Biology (2013). *The San Francisco declaration on research assessment (DORA)*. Bethesda, MD: American Society for Cell Biology. <https://sfdora.org/read>

Hicks, Diana; Wouters, Paul; Waltman, Ludo; De-Rijcke, Sarah; Rafols, Ismael (2015). "Bibliometrics: the Leiden Manifesto for research metrics". *Nature news*, v. 520, n. 7548, pp. 429. <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>

Gumpenberger, Christian; Glänzel, Wolfgang; Gorraiz, Juan (2016). "The ecstasy and the agony of the altmetric score". *Scientometrics*, v. 108, n. 2, pp. 977-982. <https://doi.org/10.1007/s11192-016-1991-5>

Mukherjee, Bhaskar; Subotić, Sinisa; Chaubey, Ajay-Kumar (2018). "And now for something completely different: the congruence of the Altmetric Attention Score's structure between different article groups". *Scientometrics*, v. 114, n. 1, pp. 253-275. <https://doi.org/10.1007/s11192-017-2559-8>

NISO (2016). *Outputs of the NISO Alternative Assessment Metrics Project*. NISO RP-25-2016. ISBN: 978 1 937522 71 1 https://groups.niso.org/apps/group_public/download.php/17091/NISO%20RP-25-2016%20Outputs%20of%20the%20NISO%20Alternative%20Assessment%20Project.pdf

Ortega, José-Luis (2018). "Reliability and accuracy of altmetric providers: A comparison among Altmetric.com, PlumX and Crossref Event Data". *Scientometrics*, v. 116, n. 3, pp. 2123-2138 <https://doi.org/10.1007/s11192-018-2838-z>

iralis®
International Registry for Authors:
Links to Identify Scientists

es:

- una **guía** para los autores hispanos para que firmen sus trabajos en el formato internacional usual
- una **base de datos** que registra las variantes de firma usadas por cada autor en diferentes épocas
- un **buscador** que usa automáticamente todas las variantes registradas

¡Ni un día más sin decidir tu firma!
Inscríbete en <http://iralis.org>

El formato de firma internacional

Con el apoyo de:

El profesional de la **información**