Relationship of the disruption indicator with other bibliometric indicators

Vicente P. Guerrero-Bote; Félix De-Moya-Anegón

Nota: Este artículo se puede leer en español en:
https://infonomy.scimagoepi.com/index.php/infonomy/article/view/33/60

Recommended citation:

Guerrero-Bote, Vicente P.; **De-Moya-Anegón, Félix** (2024). "Relationship of the disruption indicator with other bibliometric indicators". *Infonomy,* 2(1) e24016.
https://doi.org/10.3145/infonomy.24.016

Vicente P. Guerrero-Bote
https://orcid.org/0000-0003-4821-9768
https://www.directorioexit.info/ficha596
University of Extremadura
Information and Communication Department
Plazuela Ibn Marwam, s/n
06071 Badajoz, Spain
guerrero@unex.es

Félix De-Moya-Anegón
https://orcid.org/0000-0002-0255-8628
https://www.directorioexit.info/ficha92
SCImago Research Group, Spain
felix.moya@scimago.es

Abstract

An indicator to measure disruption has recently been proposed (Funk & Owen-Smith, 2017; Wu, Wang, & Evans, 2019) which has given rise to a large number of variants (Bornmann et al., 2020). In this work we are going to focus on the original indicator *DI* and the one that seems to have a better performance DI_5 (Bornmann and Tekles, 2021; Bittmann et al., 2021) carrying out a large-scale study comparing the scores assigned to each paper with other bibliometric indicators. The result is that the papers to which the bibliometric indicators assign more value do not obtain better scores. Reviews and short surveys have higher scores than articles and conference papers. Excellent papers have worse scores than non-excellent ones. Works with international collaboration obtain worse values than those without it. Works published in Q1 journals have worse

scores than those published in journals of other quartiles. And there is also a small negative correlation with the normalized impact and with the technological impact.

Keywords
Scientometrics; Bibliometric indicators; Disruption indices; scientific impact; Excellence; Technological impact.

1. Introduction

Some theories of scientific and technological change (**Schumpeter**, 2011; **Arthur**, 2007; **Tushman**; **Anderson**, 1986) consider that there are two types of advances. One type is continuity, consolidating existing knowledge streams. And another type that alters existing knowledge, making it obsolete, forcing existing theories to change. We all know both examples of consolidating advances and also disruptive advances such as the theory of relativity or the discovery of the structure of DNA.

Some studies claim that although scientific knowledge increases at an increasing rate, the disruptive knowledge generated decreases over time (**Park** *et al.*, 2023). Likewise, although many studies indicate that scientific collaboration increases the impact of results, it seems that large research teams produce more consolidating knowledge while small ones produce more disruptive knowledge (**Wu** *et al.*, 2019). To broadly study the phenomenon, they used the disruption index defined by **Funk** and **Owen-Smith** (2017). This index is based on the idea that when a focal work generates consolidating knowledge, subsequent research that cites it also includes references to the works cited by the focal work. While, if the knowledge generated is disruptive, the referenced ideas that were used for its generation lose strength and relevance, so they will not be cited by subsequent works that cite the focal work.

This disruption index has had quite an impact on the scientific literature, although they are not the only ones that have been published (**Bornmann** *et al.*, 2020; **Bornmann** and **Tekles**, 2021), variants have been proposed (**Bornmann** *et al.*, 2020; **Leydesdorff** *et al.*, 2021), versions of the indicator adapted to a specific discipline have even been developed (**Bornmann**; **Tekles**, 2019a; **Bornmann** *et al.*, 2019).

Bornmann and **Tekles** (2019b) studied the dependence of the disruption index on the citation window, concluding that after three years it stabilized.

On the other hand, the scores of different disruption indicators have also been studied. **Bornmann** and **Tekles** (2021) and **Bittmann** *et al.* (2021) conclude that one of the disruption indicators that works most consistently is the so-called DI_5 defined by **Bornmann.** *et al.* (2020).

However, there is no large-scale study comparing the scores of these disruption indicators with other more traditional bibliometric indicators. In this work our objective is to relate disruption scores with other bibliometric indicators to answer questions such as:

- Are the types of documents with the greatest scientific contribution, such as articles and conference papers, the ones that obtain the highest disruption score?

- Do excellent documents generate more disruption?
- Do good practices such as international collaboration or publication in Q1 journals have an influence on disruption?
- Is there a difference in disruption in open access publishing?
- Does disruption correlate with normalized impact or technological impact?

2. Method and data

For this work we have used all the citable documents collected in *Scopus* from 2003 to 2022 (downloaded in April 2023), which makes a total of 48,656,080 citable documents with their 1,600,096,176 references and 864,940,841 citations. As citable documents we have considered those classified by *Scopus* as: *articles*, *reviews*, *conference papers* and *short surveys.*

Bornmann *et al.* (2020) define generically for a focal work:

$$DI_l = \frac{N_i - N_j^l}{N_i + N_j^l + N_k}$$

Where:

- N_i is the number of works that cite the focal document and do not cite any of its references.
- N_j^l is the number of works that cite the focal document and at least l of its references.
- N_k is the number of works that cite one of the references of the focal document and do not cite it.

So DI_1 coincides with the disruption index defined by **Funk** and **Owen-Smith** (2017) and **Wu** *et al.* (2019) (which in this work we have simply called *DI).* And DI_5 is the one with the best behavior (**Bornmann**; **Tekles**, 2021; **Bittmann** *et al.*, 2021).

In this work, when we talk about:

- Number of items, we will be referring to the number of citable documents registered in *Scopus.*
- Excellence10, documents that are in the 10% most cited of the same year, type and category (**Bornmann** *et al.*, 2012).
- Excellence1, documents that are in the 1% most cited of the same year, type and category.
- International collaboration, documents made by authors from several countries.
- Q1, documents published in journals that are in the first quartile of their category in the *SJR* (*https://www.scimagojr.com*).
- OA, documents published in open access (labeled as such in *Scopus).*
- Normalized Impact (NI), average of the normalized citation received for each document, understood as the ratio between the citation received by the document and the average citation of documents of the same type, year and category (**Rehn**; **Kronman**, 2008).
- Technological Impact (TI), average of the normalized patent citation received for each document, understood as the ratio between the citation re-

ceived by the document and the average citation of the documents of the same year. And the citations being weighted by the GDP ratio of the countries in which the citing patents request protection (**Guerrero-Bote** *et al.*, 2021).

3. Results

The first part of Figure 1 shows the temporal evolution of both the number of items (citable *Scopus documents) and the average of the DI* and *DI*$_5$ disruption indicators. You can check the coincidence with what **Park** *et al.* (2023) reports regarding the fact that disruption indicators decrease over time.

The second and third parts of the figure show the evolution of the average disruption, but by number of authors. That is, the lines labeled 1 correspond to the evolution of the average disruption of works with one author, those labeled 2 correspond to the evolution of the average disruption of works with two authors, and so on. These second and third parts show that the disruption indicators decrease over time, but they also decrease with the number of team members, which coincides with what was reported by **Wu.** *et al.* (2019) regarding the fact that small teams generate more disruptive knowledge.

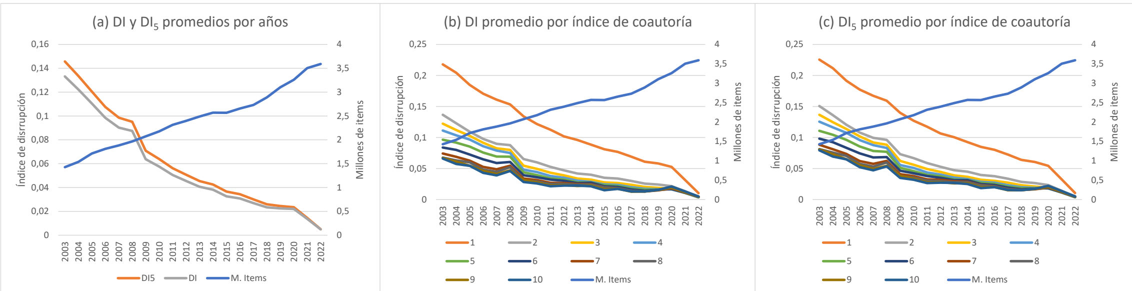

Figure 1. Temporal evolution of the number of items and the disruption indicators *DI* and *DI*$_5$ with the different co-authorship indices.

Bornmann and **Tekles** (2019b) indicated, at least a three-year citation window is needed. In this way we only consider the data until 2019 reliable.

DI and *DI*$_5$ disruption indicators by document types. Only articles, reviews, *conference papers* and *short surveys* are shown because in the study we have only taken these types of documents into account since they are the types of citable documents. Here we get the first surprise, far from what we would expect, it is not the articles and conference papers that have the highest values in the disruption indicators, but rather the reviews and short surveys, depending on the years. This is a bit counterintuitive, because among the citable documents, these types are the ones to which the least scientific contribution is attributed.

Far from what we would expect, it is not the articles and conference papers that have the highest values in the disruption indicators, but rather the reviews and short surveys, depending on the years. This is a bit counterintuitive, because among the citable documents, these types are the ones to which the least scientific contribution is attributed

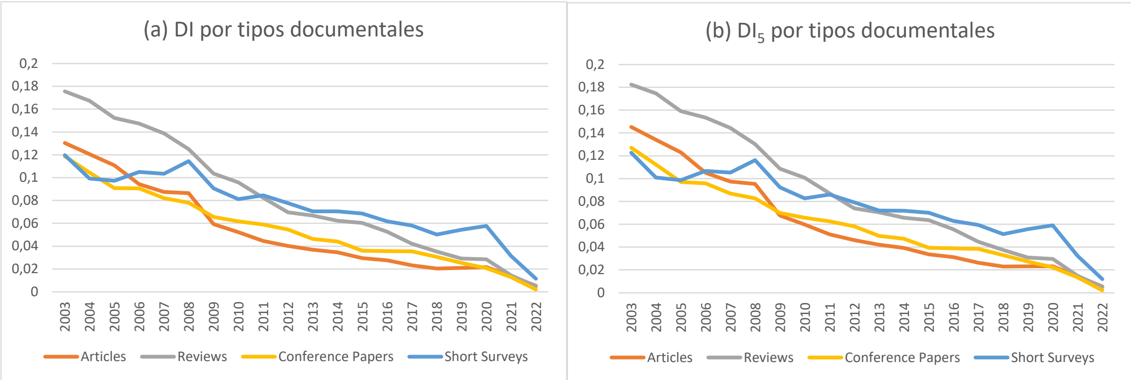

Figure 2. Temporal evolution of the *DI* and *DI*5 disruption indicators with the different types of documents.

Figure 3 shows the evolution of the average of the two indicators under study together with excellence, both excellence at 10% and excellence at 1%. In the two graphs five lines are shown, the one corresponding to the excellent documents at 10%, the one corresponding to the non-excellent ones at 10%, the one corresponding to the excellent ones at 1%, the one corresponding to the non-excellent ones at 1% and in orange that corresponding to all documents. Although the latter cannot be seen because it practically coincides with the one corresponding to the non-excellent work at 1%, which logically amounts to 99%. Above this line are only 10% non-excellent jobs, which we could say are 90% worse. And below are the excellent ones at both 1% and 10%. In the case of *DI*, excellence at 10% and 1% practically coincide, while in *DI*5 there is a small difference in favor of excellent at 1%.

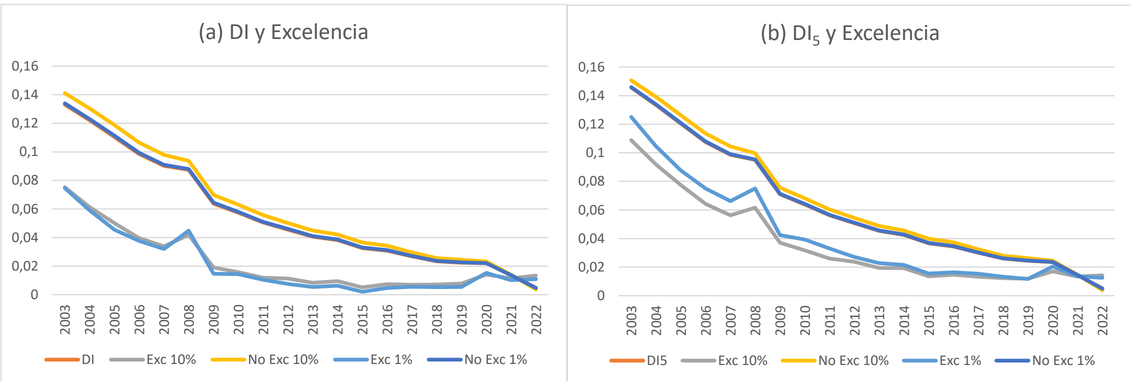

Figure 3. Temporal evolution of the *DI* and *DI*5 disruption indicators with excellence at 10% and 1%.

We can see the evolution of the averages of the indicators with international collaboration and with publication in Q1 journals in figure 4. It represents five lines, the average of the indicators, the average of works with international collaboration, the average of works without international collaboration, the average of works published in journals of the first quartile and the average of works published in journals of the second, third and fourth quartile. As can be seen in the figure, above the global figure, more disruptive knowledge is generated by works that do not have international collaboration and those that are not published in Q1 journals. This was not expected either.

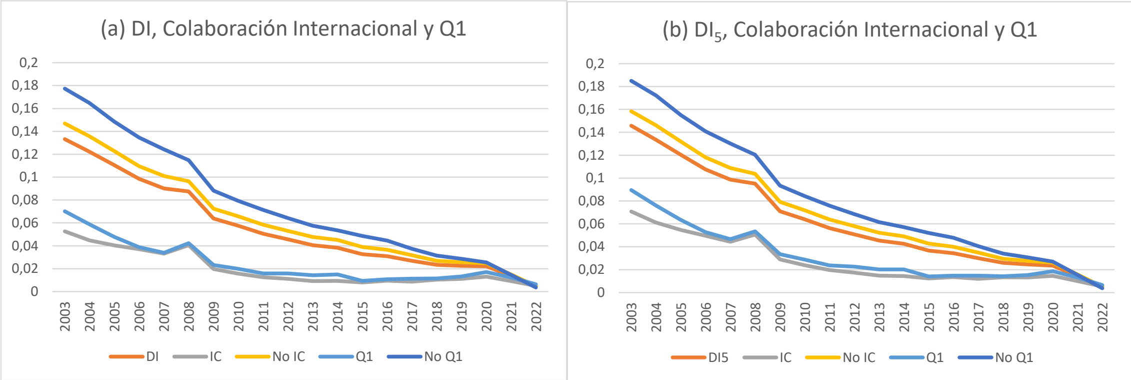

Figure 4. Temporal evolution of the *DI* and DI_5 disruption indicators with international collaboration (IC) and publication in Q1 journals.

Figure 5 shows the evolution compared to open access. As the figure shows, except in 2019 and 2020, it seems that non-open access works generate more disruptive knowledge.

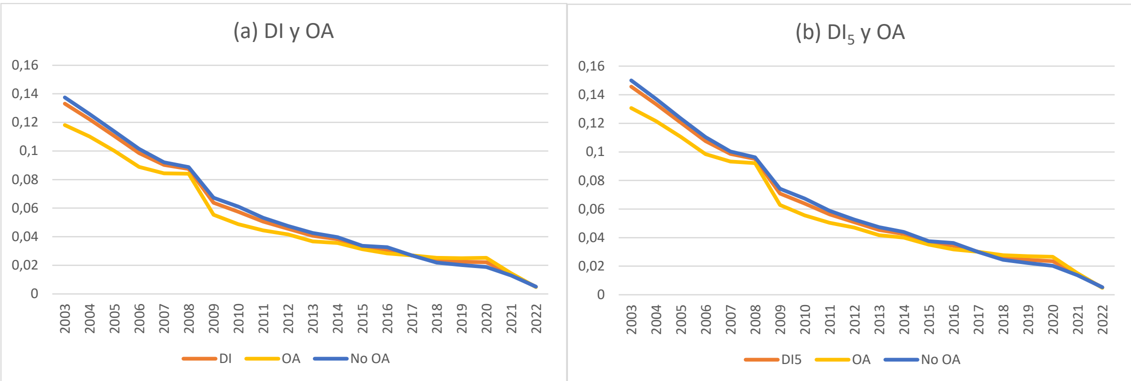

Figure 5. Temporal evolution of the *DI* and DI_5 disruption indicators with open access (OA).

Table 1 shows the correlations between the disruption indicators and the normalized impact on the one hand and with the technological impact on the other. Except in recent years, all correlations are negative, very low, but significant due to the large number of works on which they are based (except for the one with a value of -0.001, which is not statistically significant).

Table 1. Correlations of the disruption indicators *DI* and DI_5 with the normalized impact (NI) and technological impact (TI) indicators.

	***DI* -NI**	**DI_5 -NI**	***DI* -TI**	**DI_5 -TI**
2003	-0.028	-0.010	-0.011	-0.008
2004	-0.028	-0.011	-0.011	-0.007
2005	-0.024	-0.010	-0.013	-0.009
2006	-0.033	-0.017	-0.011	-0.007
2007	-0.032	-0.017	-0.009	-0.005
2008	-0.026	-0.012	-0.006	-0.003
2009	-0.027	-0.013	-0.008	-0.004
2010	-0.024	-0.011	-0.007	-0.004
2011	-0.025	-0.012	-0.006	-0.003

2012	-0.015	-0.007	-0.004	-0.001
2013	-0.024	-0.013	-0.009	-0.006
2014	-0.018	-0.010	-0.010	-0.007
2015	-0.017	-0.010	-0.008	-0.006
2016	-0.017	-0.010	-0.004	-0.003
2017	-0.016	-0.009	-0.002	-0.002
2018	-0.015	-0.007	-0.005	-0.004
2019	-0.010	-0.004	-0.007	-0.006
2020	0.002	0.007	-0.004	-0.003
2021	0.012	0.016	0.003	0.003
2022	0.043	0.048		
Total	-0.021	-0.010	-0.006	-0.004

There are no correlations with the technological impact of 2022, because that year is still too recent to be able to calculate the technological impact.

Total correlations have been made with all items from 2003 to 2019, respecting the three-year citation window indicated by **Bornmann** and **Tekles** (2019b) so that the disruption indicator is stabilized.

4. Conclusions

In this work we confirm the results obtained by **Park** *et al.* (2023) regarding the fact that the average of the disruption indicator decreases over the years and by **Wu** *et al.* (2019) that small research teams generate documents that obtain a higher average disruption score. And this confirmation is done with a different data source (*Scopus)*.

Surprisingly, the types of citable documents that have a higher average in the disruption indicators are not articles and conference papers, which are the ones that are supposed to have the most scientific contribution, but are reviews and short surveys.

Excellent documents obtain a lower average in the disruption indicators than those that are not considered excellent.

Documents produced in international collaboration obtain a lower average score in the disruption indicators. The same happens with works published in Q1 journals.

Works published in open access get slightly less disruption.

There is a very small, although mostly significant, negative correlation between the disruption indicators and the normalized impact, and also with the technological impact.

There are no major differences between the two disruption indicators studied.

Excellent documents, those produced in international collaboration, and those published in Q1 journals obtain a lower average score in the disruption indicators

More research is needed on them, because it is not credible that they correlate negatively with all the indicators used in scientometrics and bibliometrics.

5. References

Arthur, W. B. (2007). The structure of invention. *Research policy*, 36(2), 274-287.
https://doi.org/10.1016/j.respol.2006.11.005

Bittmann, F.; **Tekles, A.**; **Bornmann, L.** (2021). Applied usage and performance of statistical matching in bibliometrics: The comparison of milestone and regular papers with multiple measurements of disruptiveness as an empirical example. *Quantitative Science Studies*, 2(4), 1246-1270.
https://doi.org/10.1162/qss_a_00158

Bornmann, L.; **Devarakonda, S.**; **Tekles, A.**; **Chacko, G.** (2019). Disruptive papers published in *Scientometrics*: meaningful results by using an improved variant of the disruption index originally proposed by Wu, Wang, and Evans (2019). *Scientometrics*, 123, 1149–1155 (2020).
https://doi.org/10.1007/s11192-020-03406-8

Bornmann, L.; **Devarakonda, S.**; **Tekles, A.**; **Chacko, G.** (2020). Are disruption index indicators convergently valid? The comparison of several indicator variants with assessments by peers. *Quantitative Science Studies*, 1(3), 1242-1259.
https://doi.org/10.1162/qss_a_00068

Bornmann, Lutz; **Moya-Anegón, Félix**; **Leydesdorff, Loet** (2012). "The new excellence indicator in the World Report of the SCImago Institutions Rankings 2011". *Journal of Informetrics*, 6(2), 333–335.
http://dx.doi.org/10.1016/j.joi.2011.11.006

Bornmann, L.; **Tekles, A.** (2019a). Disruptive papers published in Scientometrics. *Scientometrics*, 120, 331–336.
https://doi.org/10.1007/s11192-019-03113-z

Bornmann, L.; **Tekles, A.** (2019b). Disruption index depends on length of citation window. *Profesional de la información*, 28(2).
https://doi.org/10.3145/epi.2019.mar.07

Bornmann, L.; **Tekles, A.** (2021). Convergent validity of several indicators measuring disruptiveness with milestone assignments to physics papers by experts. *Journal of Informetrics*, 15(3), 101159.
https://doi.org/10.1016/j.joi.2021.101159

Funk, R. J.; **Owen-Smith, J.** (2017). A dynamic network measure of technological change. *Management science*, 63(3), 791-817.
https://doi.org/10.1287/mnsc.2015.2366

Guerrero-Bote, V. P.; **Moed, H. F.**; **Moya-Anegon, F.** (2021). New indicators of the technological impact of scientific production. *Journal of Data and Information Science*, 6(4), 36–61.
https://doi.org/10.2478/jdis-2021-0028

Leydesdorff, L.; **Tekles, A.**; **Bornmann, L.** (2021). A proposal to revise the disruption indicator. *Profesional de la información*, 30(1).
https://doi.org/10.3145/epi.2021.ene.21

Park, M.; **Leahey, E.**; **Funk, R. J.** (2023). Papers and patents are becoming less disruptive over time. *Nature*, 613(7942), 138-144.
https://doi.org/10.1038/s41586-022-05543-x

Rehn, Catharina; **Kronman, Ulf** (2008). Bibliometric handbook for Karolinska Institutet. Karolinska Institutet University Library. Version 1.05.
https://doi.org/10.13140/RG.2.1.1480.9447

Schumpeter, J. A. (2011). Capitalism, Socialism, and Democracy.

Tushman, M. L.; **Anderson, P.** (1986). Technological discontinuities and organizational environments. *Administrative science quarterly*, 31, 439-465. PART III.
https://doi.org/10.2307/2392832

Wu, L.; **Wang, D.**; **Evans, J. A.** (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744), 378-382.
https://doi.org/10.1038/s41586-019-0941-9