Scientometrics (2016) 109:813–826
DOI 10.1007/s11192-016-2048-5

Effect of high energy physics large collaborations on higher education institutions citations and rankings

Edmilson J. T. Manganote[1,2] · **Peter A. Schulz**[3] · **Carlos Henrique de Brito Cruz**[4]

Received: 11 November 2015 / Published online: 8 July 2016

Abstract We analyze the effect of High Energy Physics Large Collaboration articles, an important example of Big Science and well traceable in Web of Science, on the output and citation records at the country and institutional levels. Furthermore, the effect of these specific bibliometric data on two different university rankings, the SCIMAGO and the THE, is addressed. The results suggest that these rankings may be significantly affected by this class of output, suggesting the necessity of a discussion about methodologies differentiating them from other outputs, as well as the time range considered by the rankings.

Keywords Large collaborations · Big science · University rankings · Citation records

Introduction

Collaboration is a central issue in the very definition of Science and with clear influence on research and publication practices and patterns and their associated impact. Several types of scientific collaborations have indeed evolved along the recent history and their roles

✉ Edmilson J. T. Manganote
mangano@ifi.unicamp.br

Peter A. Schulz
peter.schulz@fca.unicamp.br

Carlos Henrique de Brito Cruz
brito@ifi.unicamp.br

1 Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, Rua Sérgio Buarque de Holanda, 777, Campinas, SP 13083-859, Brazil

2 FACAMP—Faculdades de Campinas, Estrada Municipal UNICAMP-Telebrás Km 1, s/n Cidade Universitária Zeferino Vaz, Campinas, SP 13083-970, Brazil

3 Faculdade de Ciências Aplicadas, Universidade Estadual de Campinas, R. Pedro Zaccaria, 1300, Limeira, SP 13484-350, Brazil

4 Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, Rua Sérgio Buarque de Holanda, 777, Campinas, SP 3083-859, Brazil

have been studied, noticeable in the past few decades concerning a wide range of aspects. One of the main concerns is related to the importance and value of collaboration and their assessment (Katz and Hicks 1997). Particularly relevant has been the relation between the effect of international collaboration on the impact of research results (van Raan 1998). From another point of view, the effects of the intensification of scientific collaboration, as well as correlation with impact and the application of bibliometric indicators to the problem have been considered (Persson et al. 2004). More recently, the addressing of these and other close issues, like effect of collaboration on impact (Franceschet and Costantini 2010) or collaboration modes and their effects on co-authorship (Jeong et al. 2011) continue to show a growing interest.

The present paper deals with the effect of a particular kind of collaboration—very large groups of scientists expressed in long co-authorships lists—, on impact indicators when considered together with the indicators due to other kinds of collaborations. The increasing complexity of problems tackled by science started to require already in the mid of the last century the formation of large collaborations working in large-scale research facilities. This phenomenon led to the concept of *Big Science* in opposition to *Small Science*, an enterprise of individuals or small groups and resources. While the transition to Big Science was described in a seminal book by Solla Price (1963), the concept was already discussed a few years earlier (Weinberg 1961). Nevertheless, the formation of large collaboration groups, with hundreds or thousands of scientists, became necessary to reach some scientific goals. Initially associated mainly to High Energy Physics, more recently other fields such as Earth System Research, Genomics, and Clinical Medicine display similar characteristics in the necessity of putting together hundreds or thousands of contributors (Sonnenwald 2007). Indeed, the necessity of *Big Biology* in parallel to *Small Biology* deserved attention in the recent years (Vermeulen et al. 2010).

This long standing discussion ranges from the very meaning of authorship in outputs with author lists surpassing the mark of thousand co-authors (Gallison 2003; Birnholtz 2006) to the challenges of the institutional evaluation of the huge facilities that host these collaborations, like Fermilab and CERN (Irvine and Martin 1985) (Hallonsten 2014). Along with these issues, centred on the activities in the facilities themselves, a further dimension emerged in the recent years, related to the evaluation of the performance of individual researchers, on one hand committed to large collaborations and affiliated to Physics departments, but otherwise also committed to *Small Science*, in which authorship and research organization have different characteristics, where, nevertheless, discussions concerning collaborations impacts (Lee and Bozeman 2005) are also present.

A recent report by a workgroup commissioned by the International Union of Pure and Applied Physics Commission on Particles and Fields, highlights that the necessity of the participation of a large number of scientists in order to achieve the scientific goals and the procedures used in these collaborations for assigning contributions imply "that an assessment of scientific achievements based mainly on publication lists and impact factors is no longer applicable in experimental particle physics. More factors must be included to judge the scientific merits of individual researchers in this field." (IUPAP 2008). The same report makes recommendations for procedures that High Energy Physics Large Collaborations (HEPLC) should adopt to make contributions more visible.

Frequently, the articles resulting from large collaborations tend to be intensely cited in the literature, so much so that the presence of these articles in the publication list of a department tends to affect (positively) the position of the host university in international university rankings that consider the number of citations as a proxy for research impact and quality. These effects became especially noteworthy in several cases in the past few years, since the beginning of operation of the Large Hadron Collider (LHC) at CERN. Discussions pointing out this possibility start to appear on the Science discussions related blogsphere (Ranking watch 2014; Usher 2014).

Considering all aspects mentioned above, HEPLC may impact science indicators of institutions and, depending on the degree of this impact, the output and citation figures of a country as a whole. Hence, the present issue is composed by two subjects: in this work we present an analysis of the effect of HEPLC research outputs on the overall output in Physics at two levels, first at the country level and secondly at the institutional level. At the country level we consider three countries that have shown noticeable growth in knowledge production in the beginning of this century, revealing growing influence on the global science landscape after showing only modest figures prior to the 1990s: Brazil, Spain and South Korea (Almeida and Guimarães 2013). It has to be mentioned that the most impressive growth in knowledge production is China; however, as will be commented below, the effect of HEPLC is still marginal in this case. On the other hand, the participation of HEPLC is very pronounced in Brazil. Hence, we present an analysis of this effect on the article citation profile at the lowest institutional level, namely Brazilian Physics departments. Moreover, the effect of HEPLC research outputs has also to be considered in a broader context, still at the institutional level, since a citation boost for an isolated Physics Department can bias university ranking classifications. In order to scrutinize this effect for universities in emergent economies we focus on two recent rankings: (1) the Times Higher Education World Universities Ranking, namely the BRICS and emergent economies rankings (2014 edition) and (2), the 2014 Scimago Institutional Ranking for the Ibero-American universities. The analysis, following the methodology sketched below, are based on data mining on the Web of Science (WoS) platform, using the open search tools it offers.

Methodology

The objective of the work is to capture the bibliometric effect of HEPLC on the output and citation score of different institutions within the scope of publications indexed in WoS and shed light on the possible bias effects on university rankings. The impact index chosen is the citation average delivered by the citation report of WoS for a set of records, hence a widely used indicator in university policies discussions. In our search protocol for these records sets, we first selected the country of interest, choosing afterwards a specific year. The next step is dependent on the data to be considered: at the country level or at the organization level. In order to analyze the output in Physics at the country level, the search was further refined limiting the records within the 8 subfields related to Physics (Physics, Applied; Physics, Atomic, Molecular and Chemical; Physics, Condensed Matter; Physics, Fluids and Plasmas; Physics, Mathematical; Physics, Multidisciplinary; Physics, Nuclear and Physics, Particles and Fields). At this level a citation report delivers the average citations for the field. The effect of HEPLC can be obtained by excluding the articles associated to such collaborations by means of the WoS group author(s) tool. Hence a new citation report without HEPLC associated articles is obtained. At the organization level, the same procedure can be done by further choosing a specific institution, provided that care is taken by considering possible different names for the same university.

The effect of HEPLC on the average citation per article at the level of a university can be obtained by excluding directly from a given university the HEPLC's items by means of the mentioned group author(s) tool. Limiting the data collection to a succession of single year listings made it possible to obtain the citation average directly from the citation report tool option, which is limited to sets of up to 10,000 items.

The group author(s) tool identifies an organization or institution that is credited with authorship of a source publication, replacing a list of author names if the number of author names is very large (over 400, according to the definition given by Web of Science). The tool identifies the different collaboration names which can be simply selected, delivering the corresponding source items list, where individual authors and corresponding addresses are properly listed. Nevertheless, the group author(s) tool presents some limitations: not all outputs from HEPLC are necessarily addressed by this tool, as observed by inspecting the final list of records, requiring a further exclusion by hand. Part of this difficulty comes from the diversity of acronyms in identifying the same collaboration. Nevertheless, for the sake of methodological uniformity we attained to the results delivered by using the group author(s) tool, having in mind that our results can be considered as a lower limit of the relevance of HEPLC on the bibliometric indicators analysed here.

The citations averages retrieved following the above protocol are the main indicators used throughout the paper. However, in the analysis of Scimago Institutional Ranking for the Ibero-american universities, the present citations averages are correlated to an indicator provided by Scimago, namely the Normalized Impact (NI), which is a field normalized relation between the average impact of the institution and the world average. The world average has by definition NI $=$ 1. A score of 0.8 (1.3) means the institution is cited 20 % (30 %) below (above) average (González-Pereira et al. 2010).

A further indicator, derived from the citations averages obtained in the present work is the the ratio between the citations average due to the total number of outputs and the citations average excluding HEPLC outputs, called HEPLC enhancement factor (EF):

$$EF = \frac{\sum cit_{Tot} / \sum it_{Tot}}{\sum (cit_{Tot} - cit_{HEPLC}) / \sum (it_{Tot} - it_{HEPLC})} \quad (1)$$

where it_{Tot} (cit_{Tot}) corresponds to the total number of source items (citations of these items) retrieved for a given organization, while it_{HEPLC} (cit_{HEPLC}) corresponds to the sum of source items(citations to these items) of HELPC outputs. The lowest value found is EF $=$ 1, meaning that a particular institution is not involved with HEPLC.

The choice of the university rankings was determined by the noteworthy effect, as already anticipated in the introduction, of HEPLC on Brazilian Physics departments. Hence we choose two rankings that highlight contexts in which Brazilian universities are inserted among universities from other countries. First, the THE ranking for BRICS and emerging economies, considering the 25 top universities (including the two leading Brazilian universities). Secondly, the Scimago Ranking for Ibero-American Universities in which we consider all cases with NI $>$ 1.

Two further aspects should be mentioned here. First, we consider data retrieved at the end of 2014, instead of more recent retrievals, in order to have a better benchmark for the time period considered for the latest editions of the rankings discussed here. The only exception is a more recent retrieval (March, 2015) for citation data at the country level shown in Table 2. As mentioned in the introduction, we are addressing the problem at two levels: countries as a whole and institutions. At the country level, we retrieve outputs from the period 1981–2013 and focus afterwards mainly on a recent years time span: 2008–2013. At the institutional level, we retrieve outputs from 2008 to 2012, consistent with the time span also considered by the editions of the two rankings editions analyzed here. Finally, the average citations indicators are based on full citations counting, as seems to be the case for both rankings methodologies addressed in this paper. The alternative fractional citations counting, namely normalizing the citations by the number of authors started to be discussed more intensely in the past few years in somehow different contexts (Leydesdorff and Opthof 2010) and began to be considered in some rankings, like the Leiden Ranking (Leiden CWTS Ranking 2015). Possible consequences of using fractional citation by author counting are briefly discussed at the end of the present work.

Results: effect of HEPLC on research indicators

Effect at the country level

Considering that the hypothesis that HEPLC could significantly affect the records of outputs and citations, an overview over time at the country level become the first necessary benchmark. As a contextualizing starting point, we map the outputs in Physics from different countries, aiming to perceive differences in impact of HEPLC at the country level. We choose three emergent scientific countries, with some common characteristics, albeit being inserted in distinct global contexts: Brazil, South Korea and Spain. Within the common characteristics is the fact the, excluding China, these three countries present the highest outputs among the leaders in scientific output growth ranking in last years (Almeida and Guimarães 2013). Furthermore, these three countries are object of comparative studies in the literature (Fink et al. 2014); (Rodrigues and Abadal 2014).

In Fig. 1 we show the total number of items in Physics appearing in the WoS database for three different countries, South Korea, Spain and Brazil, from 1981 to 2013.

The relevance of HEPLC items can be perceived at the country level in recent years, as shown in Table 1 Having in mind the number of outputs, it should be mentioned the growing participation of HEPLC in recent years, although with different weights for different countries. As can be seen in Table 1, in 2010, considering Brazil, Spain and South Korea, only approximately 3.2, 2.2 and 1.9 % of the items are outputs from HEPLC fields, respectively. This fraction, however, increases to 7.5 % (2011) and noticeable 11 % (2012), for the case of Brazil. This increase can be attributed the participation in the LHC

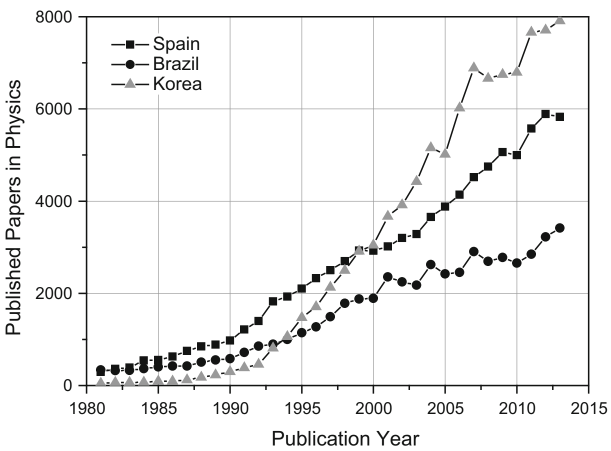


Fig. 1 Evolution of the total number of papers in Physics indexed in Wos: South Korea (*grey triangles*), Spain (*black squares*) and Brazil (*grey diamonds*)

Table 1 Total output in Physics and due HEPLC for Brazil, Spain and South Korea in recent years: 2008–2013

Year	Brazil (total)	Brazil (HEPLC)	(%)	Spain (total)	Spain HEPLC	(%)	South Korea (total)	South Korea (HEPLC)	(%)
2008	2692	51	1.9	4748	89	1.9	6661	113	1.7
2009	2780	59	2.1	5064	128	2.5	6748	133	2.0
2010	2654	85	3.2	4995	111	2.2	6793	128	1.9
2011	2847	211	7.4	5575	249	4.5	7663	205	2.7
2012	3225	365	11.3	5887	427	7.3	7708	265	3.4
2013	3416	323	9.5	5827	389	6.7	7910	260	3.3
Total	17,614	1094	6.2	32,096	1393	4.3	43,483	1104	2.5

The LHC at CERN initiated operation in 2010

collaborations, which show a boost in outputs from 2010 on (Carrazza et al. 2014). While Spain also shows a noticeable increase of HEPLC share in outputs (7 % in 2012), although not so dramatic as the Brazilian case, South Korea kept a HEPLC share of outputs in physics of only approximately 3 %. It is worth mentioning that China has not been considered here in spite of the highest growth rate of research outputs in recent years. Indeed China sums up to 367,283 in the 1981–2014 time period. However, only 3071 are due to HEPLC, a large absolute number, but corresponding to a very low share of 0.83 % of the total output. Even considering the years from 2010 onwards, this share does not surpass the 1 % rate, with a maximum of 1.3 % in 2012.

A recent and comprehensive comparison between the evolution of science and technology knowledge production in Brazil and South Korea (Fink et al. 2014) indicates a decrease in fields like physics and material science in Brazil, while South Korea shows a continuous increase in these fields. This scenario is compatible with the trends shown in Fig. 1, as well as to the rather strong relevance of HEPLC for Brazilian indicators in Physics.

A further insight of the evolution of the knowledge production in Physics in these countries can be obtained by inspecting the present citations averages of the outputs of different years, as shown for South Korea (grey diamonds), Spain (grey triangles) and Brazil (black squares), from 1981 to 2012 in Fig. 2. The data were retrieved in September 2014. For the case of Brazil, a citations average purging HEPLC records is also depicted (light grey squares). A similar data presentation has been delivered recently, focusing not on specific countries, but on publications by the American Physical Society in different subfields of Physics (Radicchi and Castellano 2011). As expected, all three countries have in common that the citations averages start to decay for outputs published after a range between 2000 and 2005, since recent publications are not yet at the end of their citation cycles (Waltam et al. 2011).

The citations average for Physics in Spain increases significantly, when comparing outputs from the early 1980s with those from late 1990s, revealing a significant higher impact, in terms of citations average than items with Brazilian and South Korean addresses. Nevertheless, the focus here is the impact of HEPLC on citations average and we chose to look closer at this impact on the output from Brazil.

As can be seen in Fig. 2, excluding HEPLC outputs (light grey squares in Fig. 2) results in noticeable decay in citations averages in the case of Brazil for recent years, but detectable since mid 1990s. Considering only the very recent records from 2012, a close

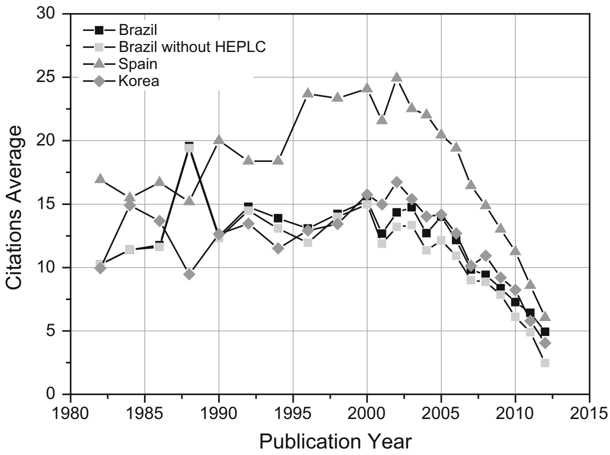


Fig. 2 Average number of citations per published item *versus* publication year for the output in Physics: South Korea (*grey diamonds*), Spain (*grey triangles*), and Brazil (*dark squares*: total production; *light grey squares*: not including HEPLC)

look to Fig. 2 reveals that the citations average of the Brazilian output drops by 50 %, if HEPLC are excluded. On the other hand, the drop in average citation with the exclusion of HEPLC outputs is 25 % for Spain and only 17 % for South Korea (not shown in Fig. 2), considering the same time window.

A general overview of the three countries can be summarized in Table 2, considering the output for a more recent time span: 2010–2014. Even including outputs at the very beginning of the LHC publication/citation cycle in 2010 (Carrazza et al. 2014), the effect of HEPLC for the three countries is noticeable, in particular for Brazil. (the numbers in parentheses represent the share in percentage of total outputs and citations due to HEPLC items in WoS).

It should be mentioned that, previously to the present work, a similar scenario, i.e., relevant positive impact of HEPLC, has been proposed, although restricted to the case of Mexico (Collazo-Reyes et al. 2010) and only at the country level.

Impact at the institutional level: Brazilian Physics departments

In view of the results of the previous paragraphs, a second step is to refine the discussion to the level of Physics departments. Here we choose to look at the leading Physics departments in Brazil, the country for which the participation of HEPLC outputs have a noteworthy weight on both share in total outputs and citations count. Here we consider the Brazilian Physics departments with the 11 highest outputs scores in 2010, data retrieved in September 2014. The most productive department presents 602 outputs, while the 11th

Table 2 Total output and citations in Physics for Brazil, Spain and Korea in the period 2010–2014 compared to the HEPLC outputs and citations in the same period and countries, as retrieved from WoS

Country (2010–2014)	Total output	Total outputs citations	HEPLC outputs	HEPLC outputs citations
Brazil	15,505	1131,441	1289 (8.3 %)	43,670 (33.2 %)
Spain	27,784	3337,12	1505 (5.4 %)	44,458 (13.2 %)
South Korea	36,763	339,255	1094 (2.9 %)	31,053 (9.2 %)

In parentheses are the percentages of HEPLC figures respective to the total outputs and citations

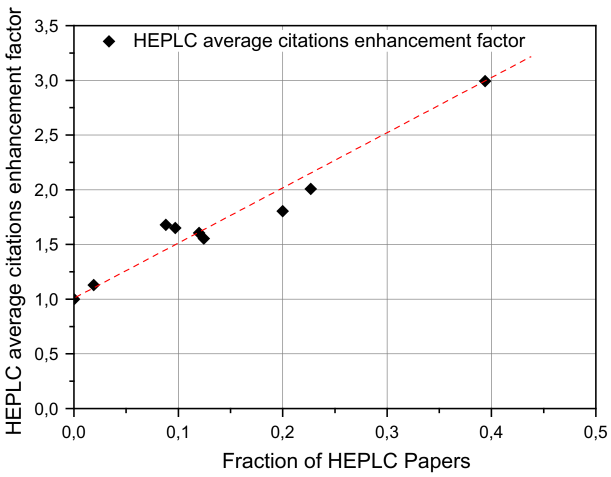
Table 3 Total and HEPLC outputs from 2010 for the 11 most productive Physics departments in Brazil at this year and the citations related to these outputs retrieved in September 2014

Institution	Total outputs	HEPLC outputs	Citations (total outputs)	Citations (HEPLC outputs)
1	602	53	6297	2877
2	267	32	3564	1609
3	233	29	2276	993
4	206	20	2666	1207
5	190	38	2761	1537
6	130	0	1693	0
7	119	27	2464	1515
8	108	0	1217	0
9	107	0	689	0
10	106	2	687	90
11	99	39	1600	1276

scored 99 outputs. The total and HEPLC related numbers of outputs, as well as the citations associated to those outputs for the 11 departments considered can be appreciated in Table 3 below.

The total number of citations and hence the citation average per item are from the WoS citation report tool. Afterwards, by excluding the HEPLC related outputs for each department, a purged citation average could be obtained. Figure 3 depicts the enhancement factor as a function of the number of outputs related to HEPLC divided by the total output of the department. Three departments showed no HEPLC related outputs and, hence, appear as a single point at the origin of the plot. The other points clearly indicate a strong relation between increasing fraction of HEPLC related outputs and the enhancement of citations averages. This is somehow expected, since those papers are outputs from a relatively low number of large collaborations, although involving a great number of institutions, are published in a core of few journals and being highly cited in a very short time range.

Fig. 3 Enhancement of citation average of leading Brazilian Physics departments as a function of the fraction of HEPLC outputs. The *vertical axis* indicate the enhancement factor: the value 2 means HEPLC outputs double the citation average. The *horizontal axis* represents the fraction of HEPLC related outputs respective to the total output of the department



The important point here is that for a significant number of departments there is a strong effect on the citations per article due to the HEPLC articles. This brings the possibility that a department assessment using bare bibliometric indicators may be biased by the participation of a few faculties active in HEPLC research collaborations. Therefore, special attention should be paid not only to individual evaluation, regarding HEPLC (IUPAP 2008), but also to the departmental assessments as a whole. Considering the enhancements shown for several Departments, it seems that the issue of fractional citations counting could receive more attention in the research assessment debate (Leydesdorff and Opthof 2010). This last point is demonstrated in more detail in the next subsection.

HEPLC impact on university rankings

We investigate the possible influence of HEPLC on two university rankings, one concerning only research performance, namely the Scimago Institutional Ranking (SIR) for Iberoamerican universities (SIR-Iber 2014); and the Times Higher Education ranking for BRICS and Emerging Economies (THE 2014), which is a "league table" type ranking. In both cases we consider de 2014 editions.

SIR is a ranking of research institutions, launched in 2009 that primarily classifies institutions by the number of outputs. Progressively SIR added more indicators in successive editions (Scimago 2012). Here we focus on the normalized impact (NI) indicator, as defined in the methodology. The NI impact considered here is obtained from the outputs within a 5 year time range, hence the 2014 edition consider the outputs in the period 2008–2012.

In Fig. 4 the citation averages of outputs from Iberoamerican universities with NI equal or above world average, according to the 2014 SIR-Iber, are shown in correspondence to the respective citations averages. The citation averages are obtained from the outputs in WoS related to each one of those universities in the same period considered in the ranking, 2008–2012. Although NI is based on citation averages, a necessary field dependent normalization is considered (Scimago 2014), while our citation averages are calculated from bare data. Furthermore, there are differences between the Scopus database, which give support to the Scimago results, and the WoS (Bar-Ilan 2010), used in the present work.

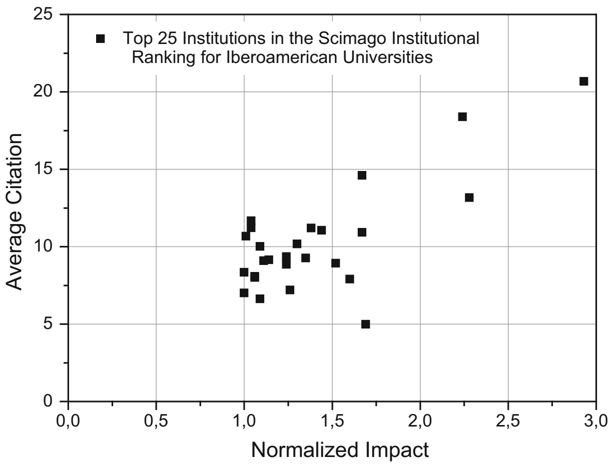


Fig. 4 Citation average obtained from WoS as a function of normalized impact (*NI*) reported in the Scimago Institutional ranking for Iberoamerican universities. Only institutions with NI equal or higher then world average and with HEPLC outputs were considered

Hence, we are not willing to, neither recalculate a NI excluding HEPLC contributions, nor quantitatively establish how Scimagós NI would change based on the variation in citation averages due to HEPLC. The point here is to call attention to trends. Indeed, the data on Fig. 4 suggests a correlation between our bare citations averages and the NI impacts from Scimago. Due to the aforementioned reasons a very strong correlation is not expected, but the correlation is rather significant, with a Pearson coefficient of 0.74 a Spearman coefficient of 0.36.

In Fig. 5 we present the citation averages excluding the HEPLC articles, for the same institutions considered in Fig. 4, as a function of the averages considering the total output, i.e., including HEPLC. The data along the dashed line represent institutions with no participation in HEPLC. Considering this framework, it is interesting to notice how the exclusion of HEPLC output leads to an important drop of citation averages, in some cases by a considerable factor of 2 and in one case by a really impressive factor of 4. It should also be noticed that, having in mind the correlation between citations averages and NI in Fig. 4, full counting of HEPLC citations could also have a noticeable impact also on the Scimagós NI, strongly suggesting that the effect of such large collaborations should be further investigated.

Another ranking, the THE ranking for BRICS and Emerging Economies, has been taken under scrutiny, revealing additional features to the relevance of HEPLC in this context. The THE is a multi indicator ranking, but the citations to papers published in the period 2008–2012, now considering the WoS database, have a considerable weight of 32.5 % (THE, 2014). As in the previous case, we are not proposing to recalculate the ranking but to call the attention to the biases that HEPLC associated items could introduce. Figure 6 shows the citations averages of the top 25 universities of the THE ranking for BRICS and Emerging Economies as a function of the NI of the same universities in the 2014 edition of the Scimago World Ranking. Here, a significant correlation (Spearmańs coefficient of 0.38) between both indicators is also present, as in the previous case shown in Fig. 4.

Figure 7 depicts how the HEPLC enhances the average citation indicator of a significant group of universities in top 25 of the THE ranking for BRICS and Emerging Economies, considering the 2012 outputs in WoS (this records are within the time span considered for the THE ranking). The citation data for the present work have been retrieved in September

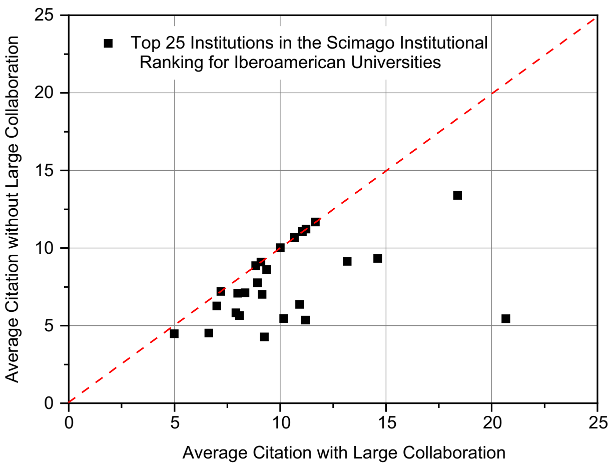


Fig. 5 Citation averages excluding HEPLC as function of citations averages for the total output of the institutions with NI $\geq$ 1 in the Scimago Institutional ranking for Iberoamerican universities also considered in Fig. 4

Fig. 6 Citation average obtained from WoS as a function of normalized impact (*NI*) reported in the Scimago Institutional Ranking for the top 25 universities in the THE BRICS and Emerging Economies ranking, 2014 edition

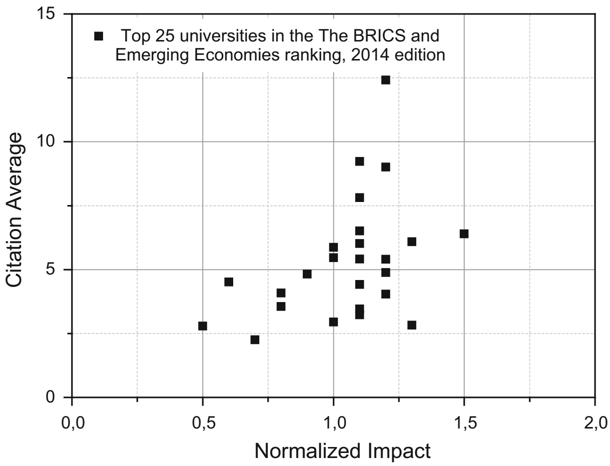
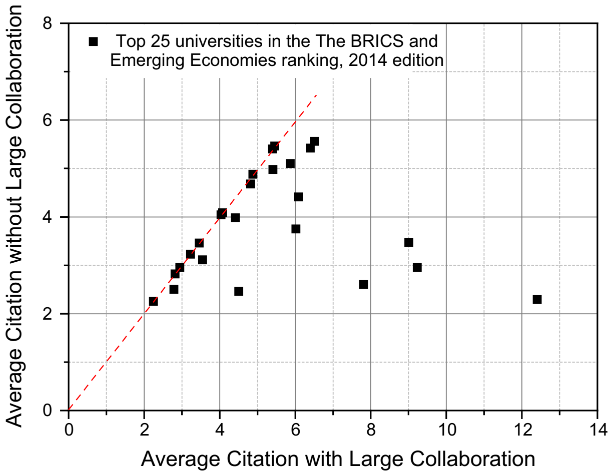


Fig. 7 Citations averages excluding HEPLC as function of citations averages for the total output of the top 25 universities in the THE BRICS and Emerging Economies ranking, 2014 edition considered in Fig. 6. The data shown here consider the outputs in WoS during 2012



2014 in order to avoid possible distortions in considering more recent data evidently not considered in the ranking. Data along the dashed line are for universities with no participation or at least a marginal one in HEPLC. The behavior seen in Fig. 7 for a set of institutions is analogous to the one in Fig. 5 for a different set of universities. It can be seen of how HEPLC may boost universities in a ranking by looking at the extreme point in Fig. 7: this point corresponds to a top 25 university, in the ranking analyzed here, for which the HEPLC outputs enhance the citations average by a factor of 6.

As already mentioned above, the set of outputs related to HEPLC is quite robust, corresponding to a rather few number of collaborations but which may include hundreds of universities addresses around the world, the outputs are published in a small core of journals and are highly and rapidly cited. The clear trend shown for the 2012 data in Fig. 7 support this claim. If we consider data from 2009 for the same top 25 universities, HELPC were much less relevant (not shown here). The influence of setting in operation the LHC by

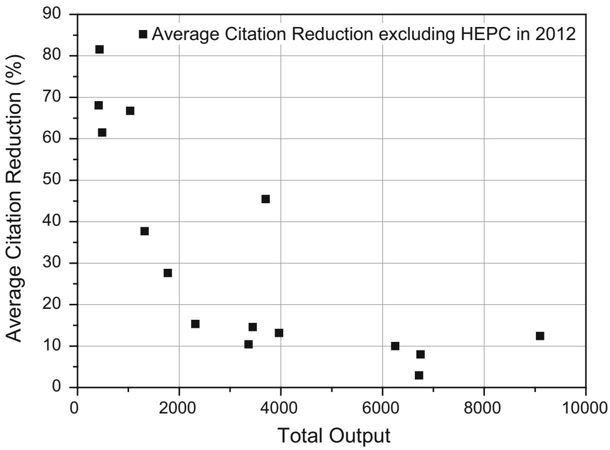


Fig. 8 Reduction in the citation average due to the exclusion of HEPLC items for 15 of the top 25 universities in the THE BRICS and Emerging Economies ranking, 2014 edition, as a function of the universities total outputs. The data shown here consider the outputs in WoS during 2012

2009 on this remarkable change has to be further investigated. It should be noticed that considering fractional citations counting could almost totally smear out the citation enhancement due to HEPLC, having in mind the great numbers of authors and participating institutions.

A further comment on this ranking is that the enhancement of citation average of universities as a whole, due to HEPLC related outputs, is inversely proportional to the total output of the corresponding universities, as shown in Fig. 8 for the 15 universities in the top 25 of the THE ranking that have participation in large collaborations: the larger the total output the smaller the effect.

Conclusions

Firstly, the conclusions shown here should not be construed as criticism of any kind to HEPLCs activities and publications. Instead, if any criticism is to be assumed, it should be directed at the use of citation indicators to define institutional or individual researchers' rankings without further consideration of the details about the way in which the research is performed.

In a broad view, the HEPLCs have a positive effect on the quantitative output and citation score at the country level in emerging countries of the "World Science System". The present work focuses mainly on data from 2008 onwards, demonstrating clearly these effects at the institutional level. Having in mind the present results, based on citation averages, as delivered by the citation reports, an analysis tool from WoS, light is shed on the relevant effects that HEPLC articles may have on university rankings and institutional assessments undertaken by policy makers.

Methodology descriptions announced by the organizers of the rankings addressed here state the use of normalization procedures in order to avoid biases: "institutions with high levels of research activity in subjects with traditionally high citation counts do not gain an unfair advantage" (THE 2014). Nonetheless, details of these methodologies should be disclosed, since it is not clear to the informed public involved in science & technology, as well as higher education policies, if effects like the influence of HEPLC are, or are not, biasing the university assessments.

In this context, it is worth mentioning two trends. On one hand, the results show how HEPLC may affect indicators at the country level, as can be seen for the Brazilian case, where total output and citation rates in Physics are noticeable increased. This effect at the country level is pervades at the institutional level, since the majority of the leading Physics departments in the country show similar effects.

This trend shown for the Brazilian case is less pronounced in Spain and South Korea. These three counties are important examples of still peripheral countries in the "World System of Science" (Leydesdorff and Zhou 2005), but showing important advances based on distinct strategies, that are of relevance to the present discussion. Indeed, regarding Physics, all three countries show a positive inflection in the production around 1990, although South Korea shows the most impressive continuous increase, while Brazil presents a noticeable slowing down after 2003 (Schulz and Manganote 2012). The significant less pronounced effect for South Korea may be attributed to the mentioned differences in strategies, since it is observed that South Korea show a much stronger trends towards Engineering and Material Sciences, closely related to Physics subfields other than Particle Physics, than Brazil (Fink et al. 2014).

However, a negligible effect at the country level does not mean that at the institutional level HEPLC are irrelevant. This is the case of China for which HEPLC have a weight of barely 1 %, while at the institutional level these collaborations boost the citation records of 3 of a total of 7 Chinese universities in the top 25 in the THE BRICS and emerging economies ranking.

Furthermore, Big Science papers may display different time cycles, since HEPLC papers are "initially highly recognized" ones, while other citation typologies are frequent in Small Science (Vlachý 1985). This is especially relevant if recent publications have the same weight as older ones in calculating normalized impacts (Waltam et al. 2011). Considering the relevant influence that HEPLC might have for some institutions in university rankings, with an effect proportional to size of the output, we suggest that a discussion should address the meaning of NI for smaller institutions and if the threshold used by THE ("exclusion of institutions that publish fewer than 100 papers a year") is adequate in order to allow for valid comparisons among institutions.

The results discussed here add elements to reassuring the importance of discussing how citations counting should be undertaken (full or fractional counting) (Huang et al. 2011) and how the research impact indicators of universities with a low total output, but with participation in large collaborations may be significantly distorted.

References

Bar-Ilan, J. (2010). Citations to the "Introduction to Informetrics" indexed by WoS Scopus and Google Scholar. *Scientometrics, 82*(3), 495–506.

Birnholtz, J. (2006). What does it mean to be an author? The intersection of credit, contribution and collaboration in science. *Journal of the American Society for Information Science and Technology, 57*(13), 1758–1770.

Carrazza, S., Ferrara, A. & Salini, S. (2014) *Research infrastructures in the LHC era: a scientometric approach. XI Milan European Economic Workshop, Working Paper*: http://wp.demm.unimi.it/tl_files/wp/2014/DEMM-2014_12wp.pdf.

Collazo-Reyes, F., Luna-Morales, M. E., Russell, J. M., & Pérez-Angón, M. A. (2010). Enriching knowledge production patternsof Mexican physics in particles and fields. *Scientometrics, 85*, 791–802.

de Almeida, E. C. E., & Guimarães, J. A. (2013). Braziĺs growing production of scientific articles—how are we doing with review articles and other qualitative indicators? *Scientometrics, 97*(2), 287–315.

De Solla Price, D. J. (1963). *Little Science, big science*. New York: Columbia University Press.
Fink, D., Kwon, Y., Rho, J. J., & So, M. (2014). S&T knowledge production from 2000 to 2009 in two periphery countries: Brazil and South Korea. *Scientometrics, 99*(1), 37–54.
Franceschet, M., & Costantini, A. (2010). The effect of scholar collaboration on impact and quality of academic papers. *Journal of Informetrics, 4*, 540–553.
Gallison, P. (2003). The collective author. In M. Biagioli & P. Gallison (Eds.), *Scientific authorship: Credit and intelectual property in science* (pp. 325–355). New York and London: Routledge.
González-Pereira, B., Guerrero-Bote, V., & Moya-Anegón, F. (2010). A new approach to the metric of journal's scientific prestige: The SJR indicator. *Journal of Informetrics, 4*(3), 379–391.
Hallonsten, O. (2014). How expensive is Big Science? Consequences of using simple publication counts in performance assessment of large scientific facilities. *Scientometrics, 100*(2), 483–496.
Huang, M.-H., Lin, C.-S., & Chen, D.-Z. (2011). Counting methods, country rank changes, and counting inflation in the assessment of national research productivity and impact. *Journal of the American Society for Information and Technology, 62*(12), 2427–2436.
Irvine, J., & Martin, B. R. (1985). Evaluating Big Science: CERŃS past performance and future prospects. *Scientometrics, 7*(3–6), 281–308.
IUPAP C11 Document 19 (2008). *Assessment of individual achievements in large collaborations in particle physics*. Retrieved in January 2015, http://www.iupapc11.org/.
Jeong, S., Choi, J. Y., & Kim, J. (2011). The determinants of research collaboration modes: Exploring the effects of research and researcher characteristics on co-authorship. *Scientometrics, 89*, 967.
Katz, J. S., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics, 40*(3), 541–554.
Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*, 673–702.
Leiden CWTS Ranking (2015) http://www.leidenranking.com/methodology/indicators.
Leydesdorff, L., & Opthof, T. (2010). Scopús soured normalized impact per paper (SNIP) versus a journal impact factor based on fractional counting of citations. *Journal of the American Society for Information Science and Technology, 61*(11), 2365–2369.
Leydesdorff, L. M., & Zhou, P. (2005). Are the contributions of China and Korea upsetting the world system of science. *Scientometrics, 63*(3), 617–630.
Persson, O., Glanzel, W., & Danell, R. (2004). Inflationary bibliomateric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics, 60*(3), 421–432.
Radicchi, F. M., & Castellano, C. (2011). Rescaling citations of publications in physics. *Physical Review E, 83*, 046116.
Ranking Watch (2014). Retrieved January 2015, http://rankingwatch.blogspot.ca/2014/10/how-to-win-citations-and-rise-in.html.
Rodrigues, R. S., & Abadal, E. (2014). Scientific journals in Brazil and Spain: Alternative publishing models. *Journal of the Association for Information Science and Technology, 65*(10), 2145–2151.
Schulz, P. A., & Manganote, E. J. T. (2012). Revisiting country research profiles: Learning about the scientific cultures. *Scientometrics, 93*(2), 517–531.
Scimago, (2012) The new excellence indicator in the world report of the SCImago institutions rankings 2011 *Journal of Informetrics*, 6 333–335.
Scimago, (2014) Retrieved in January 2015, http://www.scimagoir.com/methodology.php.
SIR-Iber, (2014). Retrieved in January 2015, http://www.scimagoir.com/pdf/iber/SIR%20Iber%202014%20HE.pdf.
Sonnenwald, D. H. (2007). Scientific collaboration. In B. Cronin (Ed.), *Annual review of information science and technology* (Vol. 41, pp. 643–681). Medford, NJ: Information Today.
THE (2014). Retrieved in January 2015: http://www.timeshighereducation.co.uk/world-university-rankings/2015/brics-and-emerging-economies/methodology.
Usher, A. (2014) *Times higher rankings, weak methodologies and the vastly overblown "Rise of Asia"*. Retrieved in January 2015, from http://higheredstrategy.com.
van Raan, A. F. J. (1998). The influence of international collaboration on the impact of research results. *Scientometrics, 42*(3), 423–428.
Vermeulen, N., Parker, J. N., & Penders, B. (2010). Big, small or mezzo. *EMBO Reports, 11*, 420–423.
Vlachý, J. (1985). Citation Histories of Scientific Publications. *The data sources. Scientometrics, 7*(3–6), 505–528.
Waltam, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: an empirical analysis. *Scientometrics, 87*(3), 467–481.
Weinberg, A. M. (1961). Impact of large-scale science on the United States. *Science, 134*(3473), 161–164.