



Unravelling the performance of individual scholars: Use of Canonical Biplot analysis to explore the performance of scientists by academic rank and scientific field[☆]



Adrián A. Díaz-Faes^{a,*}, Rodrigo Costas^b, M. Purificación Galindo^c,
María Bordons^a

^a IFS, Spanish National Research Council (CSIC), Albasanz 26-28, 28037 Madrid, Spain

^b Centre for Science and Technology Studies (CWTS), Leiden University, PO Box 905, 2300 AX Leiden, The Netherlands

^c Statistics Department, University of Salamanca, Alfonso X El Sabio s/n, 37007 Salamanca, Spain

ARTICLE INFO

Article history:

Received 5 February 2015

Received in revised form 17 April 2015

Accepted 17 April 2015

Available online 1 September 2015

Keywords:

Canonical Biplot
Multivariate analysis
Bibliometrics
Individual-level
Academic rank

ABSTRACT

Individual research performance needs to be addressed by means of a diverse set of indicators capturing the multidimensional framework of science. In this context, Biplot methods emerge as powerful and reliable visualisation tools similar to a scatterplot but capturing the multivariate covariance structures among bibliometric indicators. In this paper, we introduce the Canonical Biplot technique to explore differences in the scientific performance of Spanish CSIC researchers, organised by field (Chemistry and Materials Science) and grouped by academic rank (research fellows and three types of full-time permanent scientists). This method enables us to build a Biplot where the groups of individuals are sorted out by the maximum discriminating power between the different indicators considered. Besides, as confidence intervals are displayed in the plot, statistical differences between groups are liable to be studied simultaneously. Since test hypotheses are sensitive to different sample size effects, sizes for some pairwise comparisons are computed. We have found two gradients: a primary gradient where scientists mainly differ in terms of age, production, number of collaborators, number of highly-cited papers and their position in the byline of the publications; and a second gradient, in which scientists with the same academic rank differ by sort of field.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Scientists form the core of any research system. For this reason, increasing our knowledge on the behaviour and performance of scientists as well as on how they may be influenced by personal characteristics (Costas, van Leeuwen, & Bordons, 2010; Bozeman & Gaughan, 2011) has become an issue of great importance. Such knowledge would be helpful to better understand the research process, inform policy, and improve the ways in which scientists are considered and evaluated in their respective countries and organisations. From a bibliometric standpoint, indicators based on publications can effectively support the assessment process of individual-scholar performance, although their use at this level is not free of limitations and problems (Radicchi & Castellano, 2013). Specifically, special caution is required in the collection of data and the

[☆] The peer review process of this paper was handled by Vincent Larivière, Associate Editor of Journal of Informetrics.

* Corresponding author. Tel.: +34 91 602 28 95.

E-mail address: adrian.arias@cchs.csic.es (A.A. Díaz-Faes).

calculation of indicators due to the difficulties concerning the correct identification of the entire production of scientists and the lower validity of statistical analyses applied to small units. On the other hand, since science has become more complex over the years and demands scientists with different skills and specialisations, there is a large number of factors which should be borne in mind when analysing individual scholars. As a result, the number of bibliometric indicators has significantly grown in recent times in an effort to capture the multidimensionality of scientific activity. Even though since the appearance of the *h*-index (Hirsch, 2005) there have also been some attempts to shrink different aspects of the research performance into just one bibliometric indicator, the prevalent current belief is that an assorted set of indicators is essential in order to capture the multidimensional nature of academic activity (Moravcsik, 1984; Martin, 1996; Costas et al., 2010; Seiler & Wohlrabe, 2013).

Regarding the research performance of scientists, different dimensions such as collaboration, productivity or impact are analysed in bibliometric studies. Thus, inter-field differences in productivity and impact measures and how they may be affected by the basic or applied nature of research have attracted considerable attention in the literature (Bales et al., 2014). Moreover, since teamwork has become a distinctive feature of modern science, the benefits of scientific collaboration in impact and productivity terms has been addressed in several studies (Abramo, D'Angelo, & Solazzi, 2011a), as well as the benefits of heterogeneous (Franceschet & Costantini, 2010) and interdisciplinary collaboration, where different participants apply their own insights in a collaborative framework giving way to new knowledge (Sonnenwald, 2007). Concerning collaboration, another factor which has received special consideration has been the position of authors in the byline of scientific publications since it is determined in many disciplines according to the role played by each author in the research (Tschardtke, Hochberg, Rand, Resh, & Kruass, 2007; Waltman, 2012; Liu & Fang, 2014). First and last positions are broadly considered among experimental sciences as the most important places. First-author positions are usually occupied by those in charge of experimental work while last-author positions are reserved for those responsible for the supervision of the research. Finally, personal, institutional and environmental factors may influence the performance of research and have been examined in different studies. In particular, the role played by personal features, such as age or gender, on productivity and research impact and their relation with career success has been explored elsewhere in the literature (Costas et al., 2010; Bozeman & Gaughan, 2011; Costas & Bordons, 2011; Abramo, Cicero, & D'Angelo, 2014).

The behaviour of scientists and its relationship with personal factors may vary between fields. Exploring inter-field differences is a matter of great concern, although it still remains a challenge for bibliometrics due to field-specific productivity and citation practices (Abramo, Cicero, & D'Angelo, 2013). Interestingly, a recent paper by Ruiz-Castillo and Costas (2014) comprising a vast set of authors shows that despite great differences in productivity, it seems that their distributions are similar across fields suggesting that a single explanation of in-field variation of scientists productivity may suffice. However, from a more theoretical perspective, Bonacorsi (2008) thoroughly proves that the nature of scientific fields and the problems they face are dissimilar between old and new science fields. In this sense, old science fields like Chemistry deal with simple phenomena of increasing difficulty or with complex phenomena whose fundamental laws are well-known. On the other hand, new fields like Materials Science are placed in a region with high rates of both uncertainty and complexity. Old and new fields may differ in growth rates, degree of diversity, and types of complementarity (resources required as inputs) and this may affect the research performance of scientists. Accordingly, the mainstreaming of these new/old typologies in the analysis of individual scientific performance may also prove relevant.

On the basis of the foregoing, since scientists operate in a complex and multidimensional environment where many factors bear on their scientific activity, a large set of variables is essential for outlining their research performance. In this context, multivariate analysis is useful to provide a reliable picture of the different aspects at play in the activity of individual scholars. In many cases, data structure and relations are far too complex to be successfully addressed through univariate or bivariate methods. For this reason, when we move from a one-dimensional space into a multidimensional one some clear *prima facie* relations become blurred, but the fact is that one-dimensional approaches are likely to provide inaccurate results since patterns are complex and cannot be described unambiguously (Moravcsik, 1984).

In this paper, we consider that given a set of scientists, grouped by any specific characteristic (e.g. age, academic rank, gender), and a large set of bibliometric indicators that describe their activity, it would be relevant to be able to find out which set of bibliometric indicators discriminate best among the different groups. For this purpose we have chosen Biplot methods since they are powerful and rich visualisation tools, similar to a scatterplot but capturing the multivariate covariance structures among bibliometric indicators. Biplot representations were originally proposed by Gabriel (1971) as methods of multivariate analysis which provide a joint plot of rows and columns in a low-dimensional Euclidean space using markers (points/vectors) for each of them chosen in such a way that the scalar product represents the elements of the data matrix. The markers are obtained by the usual *singular value decomposition* (SVD) of a matrix $X_{n \times p}$, where n usually refers to the number of rows (elements) and p represents the number of columns (variables) measured on them. Then, the matrix is factorised in row and column markers. Since there are multiple ways to factorise a data matrix, the Biplot representation will have different properties according to the metric selected. They can disclose patterns of covariation and correlation, differences between groups of sample units and, most importantly, the relation between individual units and the multivariate structure of the data (Gabriel & Odoroff, 1990). Gabriel basically described two types of Biplots (classical Biplots): JK-Biplot in which only the rows are represented with high quality (row-metric preserving), and GH-Biplot in which only the columns are represented with high quality (column-metric preserving). A comprehensive study of the different focuses and alternative types of the Biplot can be found in Cárdenas, Galindo, and Vicente-Villardón (2007). In as far as the field of bibliometrics is concerned, Biplot methods were introduced by Díaz-Faes, Benito-García, Martín-Rodero, and Vicente-Villardón (2011) and

Table 1
Number of individuals by area and rank.

	Chemistry	Materials Science	Total
Post-doc	26	36	62
Tenured scientist	160	179	339
Research scientist	85	82	167
Research professor	72	89	161
Total	343	386	729

Díaz-Faes, González-Albo, Galindo, and Bordons (2013) and their practical usefulness has been proved in a previous work which explore the performance of networking research centres (Morillo, Díaz-Faes, González-Albo, & Moreno, 2014). These papers consider the HJ-Biplot (Galindo, 1986), which takes the good properties of JK-Biplot and GH-Biplot and provides an optimum quality of representation (QLR) for both points and vectors in the same Cartesian system. Although focused on the JK-Biplot method, another recent study has also discussed its potential in the field of bibliometrics (Torres-Salinas, Robinson-García, Jiménez-Contreras, Herrera, & Delgado López-Cózar, 2013).

In this work, we aim to introduce a new version called 'Canonical Biplot' (Gabriel, 1972; Vicente-Villardón, 1992), which maintain the intrinsic features of the Biplot analysis and allows for a statistical exploration to find out whether there are significant differences in mean values between groups of authors and to identify which bibliometric indicators account for such differences. An appealing application in geology can be found in Varas, Vicente-Tavera, Molina and Vicente-Villardón (2005). Accordingly, this version provides a simultaneous representation of rows (groups) and columns (indicators) in such a way that the groupings are separated by the largest discriminating power between them. In this study, we explore individual performance from different perspectives: collaboration, research level (basic/applied), interdisciplinarity, impact, production, authorship position, academic rank, age and discipline. Finally, some research questions are addressed by means of the integrated analysis of groups and variables: which variables discriminate best between groups of individuals and research fields? Can we anticipate specific patterns or any particular role for scientists based on their academic rank? Which bibliometric indicators best define each group?

2. Methods

2.1. Data and bibliometric indicators

This study is based on the analysis of the scientific publications of 729 active scientists affiliated to the Spanish National Research Council (CSIC) in 2007 in the research areas¹ of 'Chemistry Science and Technology' [hereinafter referred to as 'Chemistry'] and 'Materials Science and Technology' [hereinafter referred to as 'Materials Science']. All active scientists included in this study are grouped according to their academic rank in 2007 as defined by the CSIC: 'Post-doc' researchers (which includes various types of research fellows), 'Tenured Scientists' (lowest tenured rank), 'Research Scientists' (intermediate tenured rank), and 'Research Professors' (highest tenured rank). Personal data such as full name, age and academic rank were provided by the CSIC. Table 1 shows a summary of the number of individuals in each different group by rank and research area.

Publications from the Web of Science (WoS) database during the period 2007–2011 were collected for all the individuals included in this study. For a proper allocation of scientific production to individual scholars, an ad hoc software designed to cope with inconsistencies regarding author's names was used in order to automatically collect the publications of the researchers included in this study. A manual revision was conducted to correct minor inconsistencies. As regards citations, this analysis focuses on articles and reviews (hereinafter referred to as "papers") and citations to these papers were collected up to 2013.

Based on all the information above, the following set of indicators was obtained for each researcher:

- *Age*: age of the researcher at the beginning of period (i.e. 2007).
- *Production*: the total number of papers published by a researcher during the entire period (2007–2011).
- *Pratt index*: measures the concentration of any given researcher papers by subject categories according to the assignment of the publication journals to WoS categories (Pratt, 1977).

$$C = \frac{2 \left(\frac{s+1}{2} - q \right)}{s - 1}$$

with

$$q = \sum \frac{ia_i}{t}$$

¹ The CSIC comprises approximately 130 research institutes grouped into eight different research areas for the sake of research management (<http://www.csic.es/web/guest/areas-cientificas>). Every CSIC scholar is assigned to one of these areas.

where s represents the number of subject categories of the publications, a_i is the size of the category of rank i and t is the total number of publications. It ranges from $C = 0$ which corresponds to a uniform distribution of papers across the categories, to $C = 1$ that corresponds to the case where all papers belong to exactly one category. The lower the concentration, the higher is the interdisciplinarity of the researcher. It is a small variant (normalisation) of the better known Gini index (Carpenter, 1979).

- *Research level (RL)*: a classification scheme of four research levels was used to describe the basic or applied orientation of a journal. Journals were assigned to a research level on the basis of both expert review and patterns of journal-to-journal citation. It ranges from 1 to 4, where 1 represents the highest level of applied research and 4 refers to the most basic level of research. This classification was described by CHI Research/Computer Horizons Inc. (Narin, Pinski, & Gee, 1976), which now operates as ipIQ. A mean value was assigned to each researcher according to his/her publication pattern.
- *G-index*: given a set of papers ranked in decreasing order by the number of citations, g is the highest (unique) number of papers that together received g^2 or more citations (Egghe, 2006). It combines a measure of quantity (articles) and impact (citations). It was proposed as an improvement of the h -index, since it discriminates better for highly skewed citation distributions.
- *First author*: proportion of papers that each researcher has published as first author.
- *Last author*: proportion of papers that each researcher has published as last author².
- *Number of collaborators*: this indicator reflects the distinct number of co-authors with whom one author has published.
- *Co-authorship index*: this indicator quantifies the average number of authors per publication of a given researcher.
- *Top10%*: this indicator measures the number of papers of a researcher which, compared with other similar papers (same field, publication year, and document type), belong to the top 10% most frequently cited, i.e. Percentile 90.
- *Mean Normalised Citation Score (MNCS)*: measures the average normalised number of citations received by the papers of a researcher.

$$MNCS = \frac{1}{n} \sum_{i=1}^n \frac{c_i}{e_i} \text{ where } n \text{ is the number of papers of a researcher, } c_i \text{ denotes the number of citations received by paper } i \text{ and } e_i \text{ represents the average number of citations of all WoS publications published in the same subject category, in the same year and that have the same document type as paper } i \text{ (Waltman, van Eck, van Leeuwen, Visser, \& van Raan, 2011). An MNCS = 1 can be interpreted as the world average, so if MNCS > 1 for a given researcher, it means that he/she has received on average more citations than the world average and vice versa.}$$

i and e_i represents the average number of citations of all WoS publications published in the same subject category, in the same year and that have the same document type as paper i (Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011). An MNCS = 1 can be interpreted as the world average, so if MNCS > 1 for a given researcher, it means that he/she has received on average more citations than the world average and vice versa.

- *Mean Normalised Journal Impact (MNJS)*: is the average normalised citation score of the journals in which a CSIC researcher has published compared to the world average in the same field, year, and document type. Like MNCS, the MNJS indicators corrects for difference among fields.
- *Proportion of national collaborative publications (pp nat collab)*: this indicator measures the proportion of papers that a researcher has carried out in collaboration between two or more Spanish organizations. It may also include international collaborative papers.
- *Proportion of international collaborative publications (pp int collab)*: the proportion of co-authored publications with authors affiliated to two or more countries.
- *Proportion of collaboration with industry (pp industry)*: this indicator quantifies the proportion of papers that a researcher has co-authored with one or more authors affiliated to an industrial partner (Tijssen, 2011).
- *Mean Geographical Collaboration Distance (MGCD)*: the geographical collaboration distance is defined as the largest geographical distance between two addresses reported in the publication's address list. If a paper contains just one address, then MGCD = 0. A detailed description of the geocoding procedure can be found in Waltman, Tijssen, and van Eck (2011).
- *Proportion of long distance collaborative publications (pp long dist collab)*: proportion of author publications that have a geographical collaboration distance ≥ 1000 km.

2.2. Canonical Biplot

Gabriel (1972) and Vicente-Villardón (1992) have proposed the Canonical Biplot as an alternative Biplot technique to other similar techniques such as MANOVA, Discriminant Analysis or Canonical Analysis. Given a data matrix $X_{n \times p}$, suppose that n rows (in our case 729 authors) can be divided into K clearly differentiated groups (e.g. eight groups in our case, this is the combination of the authors' field and rank categories) with n_k authors in each and we have measured p variables (in our case, the above mentioned 17 indicators) for each of them. Taking the matrix of means and the covariance matrices between and within the groups it is possible to build a Biplot where the groups are separated by the highest discriminating power between them. If we project the groups on the plot, the coordinates on the first axis represent the linear combination of variables that produces the largest univariate Snedecor F in the ANOVA. Accordingly, the bibliometric indicators with larger F values will show the highest differences between the groups. This provides a Biplot representation with the following properties:

² Single-authored papers are included in both, first and last author indicators, but only account for 1% of the total number of publications.

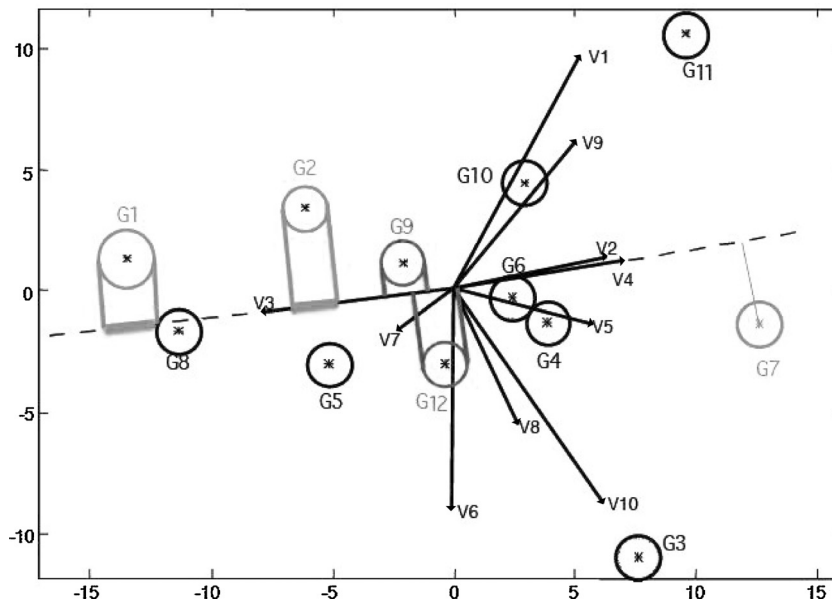


Fig. 1. Example for the interpretation of the first factorial plane of a Canonical Biplot.

(1) discriminant coordinates (projections in the directions of largest separation), (2) the Euclidean distance between two means markers approximates the Mahalanobis distance between groups (magnitude of effects) and (3) it enables us to place region predictions on the factorial plot.

To illustrate the interpretation of the Canonical Biplot analysis, we provide an example (Fig. 1). Row markers (stars) depict the average values of the different groups, while vectors represent the different variables (indicated also under V labels). If a group marker projection is close to a variable (end or prolongation of the vector), it means that the average of group is high for that indicator. Conversely, if it is far away it will only take low values. For instance, in Fig. 1 group G7 has large mean values for variables V2, V4 and V5 (as it is close to them in the projection), while the same group has low mean values for V3, for which groups G1 and G8 obtain large average values.

Besides, the angles between vectors (variables) can be interpreted as an approximation of their correlation. Thus, acute angles between two indicators (for instance, V8 and V10, or V2 and V4) indicate that such variables have a strong positive correlation, while obtuse angles (V3 and V4) represent high negative correlations, and right angles (V6 and V4) imply uncorrelated variables. As in other Biplots, we have some measures on the quality of representation and the goodness of fit for bibliometric indicators and, regarding this particular method, for the group means. As the factors obtained are uncorrelated, the inertia (variance) of a variable explained on a particular plane is the sum of the contributions to the axes that form that plane, and that quantity is known as quality of representation (QLR). It is important to bear in mind that only variables and groups with a good QLR should be interpreted on a given plane. In this study, we measure QLR on a 100-point scale.

Confidence intervals are shown along with group's markers in such a way that the projections of the circles onto the direction of a variable approximate a confidence interval. If we project two circles perpendicularly and both intervals do not overlap at all, this indicates that there are significant differences between the means of the groups ($p < 0.05$) (e.g. G1 and G2 over V3 in Fig. 1). However, should an overlap occur, as is the case between G9 and G12 over V3, we may infer that there are not significant differences for that variable. The radius of the circle is calculated as $t_{n-K,\alpha} / \sqrt{nK}$, where $t_{n-K,\alpha}$ is the critical point of a Student's t distribution with $n - K$ degrees of freedom, for a significance level of α (see Amaro, Vicente-Villardón, & Galindo, 2004 for a detailed description). As the radius of the circle is based on a Student's t distribution, a Bonferroni correction, $K = r(r - 1)/2$ where K is the number of possible comparisons and r is the number of groups, was added to correct for multiple comparisons which increase the size of the circles and reduce error Type I (wrong rejection of null hypothesis)³. Indicators were standardised by columns since they do not have the same measuring scale. The statistical analysis was performed using a Matlab program (Vicente-Villardón, 2014).

2.3. Effect sizes

Although our study did not intend to accept or reject null hypothesis for each indicator (from a more classical significant test hypothesis perspective) but to find top discriminant indicators for the different groupings, there are some intrinsic

³ Note that the method is conservative since where two circles do not overlap we may assume that there is a significant difference, but if there is an overlap we may find a significant difference in another direction of the multidimensional space (Varas et al., 2005).

limitations to the results of significance tests described in the literature that we have to bear in mind (Kirk, 1996; Schneider, 2013). In this study, Post-doc groups have a small sample size as compared to permanent scientists (i.e. Tenured Scientists, Research Scientists and Research Professors). Therefore, we may argue that significance tests are more sensitive for Post-docs in terms of wrong acceptance of the null hypothesis (Type II error). Thus, we have calculated the size effect using Hedge's g (Hedges, 1981), which is similar to Cohen's d (Cohen, 1988) but weighing each group's standard deviation by its sample size. This effect quantifies the magnitude of the differences between groups of scientists expressed in standard deviation units. Since $g \cong d$, we have taken the values proposed by Cohen (1988), which are widely accepted: <0.20 trivial, ≥ 0.20 small effect, ≥ 0.50 medium effect, ≥ 0.80 large effect and ≥ 1.30 very large effect⁴. Effect sizes were computed by means of the R package 'compute.es' (Del Re, 2013).

3. Results

CSIC scientists published a total of 9163 papers in the WoS database during the period 2007–2011 with a resulting breakdown by research area of 4886 in Materials Science (MAT) and 4277 in Chemistry (CHEM). Table 2 displays the means and standard deviations for the 17 variables analysed by academic rank and field. We provide this information to support our subsequent results.

A glimpse at Table 2 gives us some preliminary hints about the scientific performance of the researchers under study. For instance, g -index is larger for higher academic positions in both areas, ranging from $g=7$ or 8 for Post-docs to $g=13$ for Research Professors in both research areas. However, excluding Post-docs, scientists in a given rank show very similar g -index in both areas. When the research level is considered, it tends to be more basic on average for Chemistry scientists regardless the rank when compared to Materials Science. Interestingly, collaboration with industry reaches its highest point at the lowest and intermediate tenured positions (Tenured Scientists and Research Scientists) and shows low dispersion for all scientists in both the Chemistry and Materials areas. MNJS shows that Post-docs publish on average in slightly more outstanding journals in terms of citations, although their values are more scattered across the distribution. All in all, the information in Table 2 is highly revealing but also very massive, and it is just here where the use of the Canonical Biplot can help to better summarise and discuss the differences and main patterns among the different indicators and groups of researchers, as described in the section below.

3.1. Biplot analysis

The factorial plot resulting from the Canonical Biplot analysis is displayed in Fig. 2. We analyse the main factorial plane which accounts for most of the variance. The value of Wilk's lambda distribution ($\lambda = 7.914$ with $p < 0.01$) shows that the difference between the means of groups actually exists, i.e. there are indeed differences that can be attributed to the groups under survey. The inertia absorption for the first factorial plane is 91.35%. Axis 1 accounts for 80.23% of the variance whereas Axis 2 explains 11.12%.

Since Canonical Biplots provide discriminant coordinates, a straightforward separation among the groupings of scientists can be clearly seen in the first factorial plane. The most remarkable feature is that there are major discrepancies between scientists from different academic ranks since they show singular features which separate them from the rest, especially in the lowest and highest ranks. There are global differences in the ANOVA test ($p < 0.01$) for all indicators except for pp nat collab, pp industry and MNCS. The coordinates of Axis 1 show that the discrepancies between Post-docs and the three types of full-time scientists are mainly based on age, production, number of collaborators and order of signature (first author, last author). These indicators have the largest univariate Snedecor F in the ANOVA. Accordingly, the variables that best separate one group from another present high contributions to the axes and those who present low discriminatory power have a poor QLR (coloured in grey in Fig. 2) and their information is put aside by the Biplot (Table 3). For further analysis on the discrepancies between pairs of groups, we focus on the projections of the circles in the direction of the variables.

When the circles of the groups are projected onto different vectors, we observe for Axis 1 that there are statistical differences (the projections of the circles do not overlap) on age, production, g -index, number of collaborators, ptop10% and the proportion of papers published as first author for each academic rank. Thus, as we move from left to right across the plot, researchers with the highest rank (Research Professors) are older and show higher rates of performance with a growing number of papers published, their networks in terms of collaborators become larger and their rate of papers with a high number of citations (acute angle between vectors) is higher. On the other hand, as scientists move up in academic rank, interdisciplinarity (Pratt index) tends to shrink and their publications are concentrated in a smaller number of subject categories. Interestingly, the measures of individual citation performance vary according to the selected indicator. For size-dependent indicators, such as ptop10% and g -index, a better citations performance is found for higher ranks. Nevertheless, if we take the MNCS for reference, the behaviour on average is similar for all scientists as we have already pointed out (no global differences). Regarding the order of signature, it appears, for instance, that last authors seem to be more common at higher academic ranks, although there are no significant differences for some pairwise comparisons such as between Researcher Scientists and Research Professors in Chemistry (projections of circle do overlap). On the other hand, Post-docs

⁴ These thresholds are informative and easy to understand but they should not be used uncritically as benchmarks (Cohen, 1988).

Table 2
Average values for individual-level indicators.

Groups	Age	Pratt Index	Research Level	g-Index	Production	First author	Last author	N Collaborators	Co-authorship index
Post-doc (CHEM)	35 ± 5	0.33 ± 0.20	3.21 ± 0.68	7 ± 6	10 ± 10	0.37 ± 0.28	0.07 ± 0.11	24 ± 23	5.58 ± 1.27
Tenure (CHEM)	44 ± 8	0.43 ± 0.18	3.18 ± 0.75	10 ± 6	16 ± 12	0.15 ± 0.22	0.22 ± 0.24	36 ± 30	6.02 ± 1.61
Researcher (CHEM)	51 ± 7	0.45 ± 0.19	3.33 ± 0.73	11 ± 8	21 ± 25	0.09 ± 0.17	0.36 ± 0.28	40 ± 33	5.80 ± 1.56
Professor (CHEM)	59 ± 7	0.47 ± 0.20	3.32 ± .59	13 ± 12	30 ± 46	0.08 ± 0.17	0.32 ± 0.26	60 ± 79	5.60 ± 1.31
Post-doc (MAT)	35 ± 3	0.40 ± 0.16	3.03 ± 0.65	8 ± 4	11 ± 6	0.31 ± 0.22	0.06 ± 0.09	24 ± 15	5.81 ± 1.77
Tenure (MAT)	45 ± 8	0.46 ± 0.19	2.95 ± 0.61	10 ± 6	18 ± 11	0.16 ± 0.15	0.18 ± 0.15	39 ± 28	5.46 ± 1.50
Researcher (MAT)	51 ± 8	0.46 ± 0.17	3.00 ± 0.62	11 ± 6	21 ± 13	0.14 ± 0.16	0.24 ± 0.19	43 ± 31	5.31 ± 1.27
Professor (MAT)	57 ± 7	0.50 ± 0.18	3.14 ± 0.58	13 ± 10	30 ± 23	0.08 ± 0.09	0.30 ± 0.17	61 ± 52	5.58 ± 1.38
Total CHEM	48 ± 10	0.44 ± 0.19	3.3 ± 0.71	11 ± 8	20 ± 27	0.14 ± 0.22	0.27 ± 0.26	41 ± 46	5.84 ± 1.52
Total MAT	48 ± 10	0.44 ± 0.19	3 ± 0.61	11 ± 7	21 ± 16	0.15 ± 0.16	0.21 ± 0.17	44 ± 36	5.50 ± 1.46

Groups	Ptop10%	MNCS	MNJS	pp int collab	pp nat collab	pp industry	MGCD	pp long dist collab
Post-doc (CHEM)	1.38 ± 4.28	0.92 ± 0.65	1.48 ± 0.60	0.31 ± 0.25	0.56 ± 0.35	0.020 ± 0.06	918 ± 817	0.28 ± 0.24
Tenure (CHEM)	2.09 ± 2.97	1.15 ± 1.15	1.35 ± 0.39	0.33 ± 0.25	0.51 ± 0.30	0.042 ± 0.08	1592 ± 1595	0.29 ± 0.24
Researcher (CHEM)	2.37 ± 3.95	0.99 ± 0.66	1.36 ± 0.44	0.34 ± 0.26	0.54 ± 0.32	0.041 ± 0.11	1,662 ± 1,829	0.29 ± 0.25
Professor (CHEM)	5.62 ± 15.44	1.01 ± 0.73	1.32 ± 0.42	0.33 ± 0.24	0.53 ± 0.31	0.030 ± 0.05	1,640 ± 1,600	0.29 ± 0.23
Post-doc (MAT)	0.90 ± 0.94	1.05 ± 0.52	1.41 ± 0.58	0.39 ± 0.29	0.49 ± 0.34	0.024 ± 0.05	1,189 ± 1,169	0.31 ± 0.25
Tenure (MAT)	1.95 ± 2.35	1.04 ± 0.79	1.26 ± 0.54	0.42 ± 0.27	0.49 ± 0.30	0.037 ± 0.09	1,869 ± 1,595	0.36 ± 0.24
Researcher (MAT)	2.31 ± 3.31	1.00 ± 0.76	1.26 ± 0.49	0.39 ± 0.22	0.54 ± 0.28	0.034 ± 0.06	1,632 ± 1,482	0.32 ± 0.20
Professor (MAT)	3.46 ± 6.40	1.14 ± 1.52	1.32 ± 0.54	0.45 ± 0.25	0.51 ± 0.27	0.024 ± 0.04	2,003 ± 1,584	0.37 ± 0.23
Total CHEM	2.85 ± 7.8	1.07 ± 0.77	1.36 ± 0.43	33 ± 0.25	0.53 ± 0.31	0.037 ± 0.08	1,568 ± 1619	29 ± 0.24
Total MAT	2.85 ± 7.8	1.06 ± 0.98	1.29 ± 0.53	0.42 ± 0.26	0.51 ± 0.29	0.032 ± 0.07	1,786 ± 1546	35 ± 0.23

Note: Data expressed as average ± standard deviation.

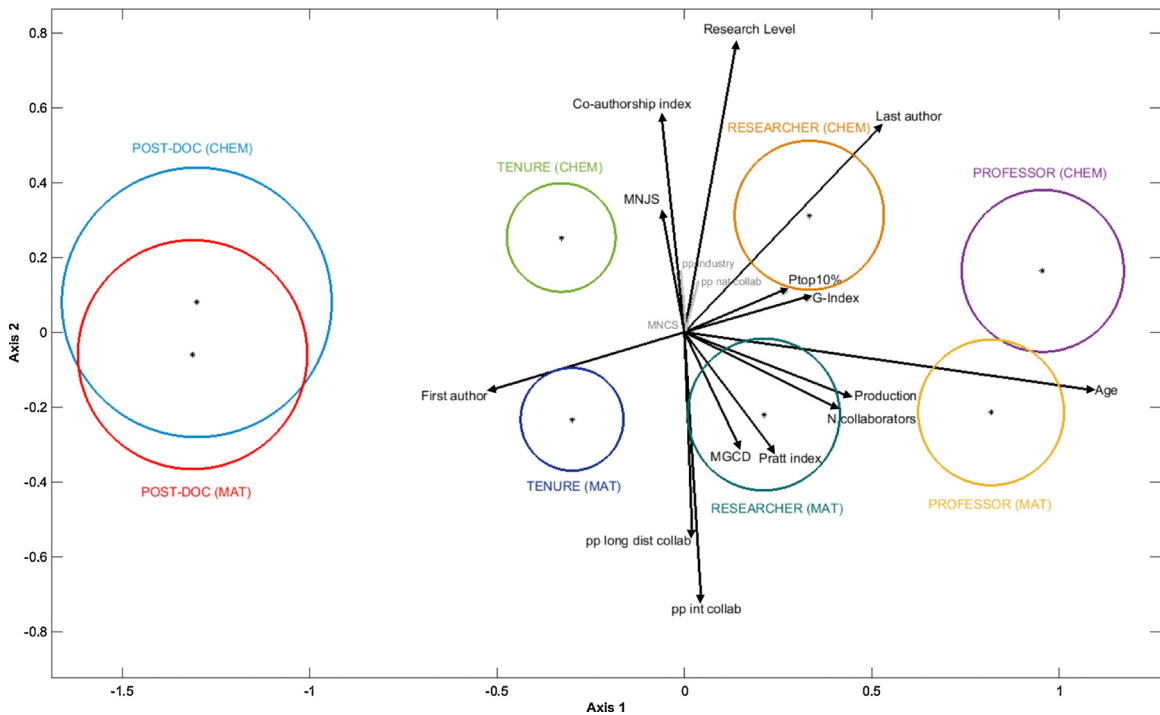


Fig. 2. Two dimensional Canonical Biplot for authors performance (plane 1–2). Variables coloured in grey present a QLR below 50 points.

most distinctive feature, which separate them from the rest of the groups, is the higher proportion of papers signed as first authors. This variable has a high negative correlation (obtuse angles) with the aforesaid indicators for which scientists with high academic rank show top values. All in all, the first position of signature seems to discriminate very well between low and high academic ranks in both fields.

Concerning Axis 2, we observe there are inherent features among full permanent scientists for each field due to the fact that scientists within the same rank perform alike in both fields (Fig. 2). Chemistry scientists are placed on top of the plot performing on average a more basic and slightly more interdisciplinary research with a higher number of collaborators (group markers are close in projection to research level and the co-authorship index and far away from the Pratt index). Materials Science researchers, which are placed at the bottom, collaborate more at international level with more distant partners and are involved in more applied research. Moreover, a higher MNJS is observed in Chemistry when compared to that in Materials Science.

On the other hand, we have also depicted plane 3–5 where most of the variance for collaboration with industry and MNCS is explained (Fig. 3). This plane can be interesting in order to observe that these two indicators show a weak discriminatory

Table 3
Quality of representation (QLR) to explain the group means.

Variables	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
Age	99.40	0.28	0.13	0.13	0.01
pp nat collab	11.48	21.25	17.59	14.17	14.08
Pratt index	65.55	16.09	13.73	0.34	2.15
Research level	16.72	72.53	7.16	2.13	0.58
g-index	96.00	1.11	0.21	0.15	1.49
Production	92.91	1.96	2.38	0.00	1.55
First author	80.25	0.99	17.6	0.77	0.07
Last author	79.35	12.14	3.00	3.47	1.52
N collaborators	88.68	3.03	1.3	1.42	5.42
Co-authorship index	5.54	69.12	5.45	0.00	19.25
Ptop10%	70.89	1.78	13.12	11.02	0.68
MNCS	0.03	1.41	27.5	4.17	60.85
MNJS	10.25	41.91	25.79	11.19	9.82
pp int collab	2.33	82.43	4.85	4.9	2.91
pp industry	1.58	19.30	58.66	6.75	13.65
MGCD	34.68	20.98	34.44	0.91	2.25
pp long dist collab	0.85	78.37	10.32	0.31	3.85

* Bold values > 50 (a good QLR is considered when at least half of the variance of an indicator is explained on a particular axis).

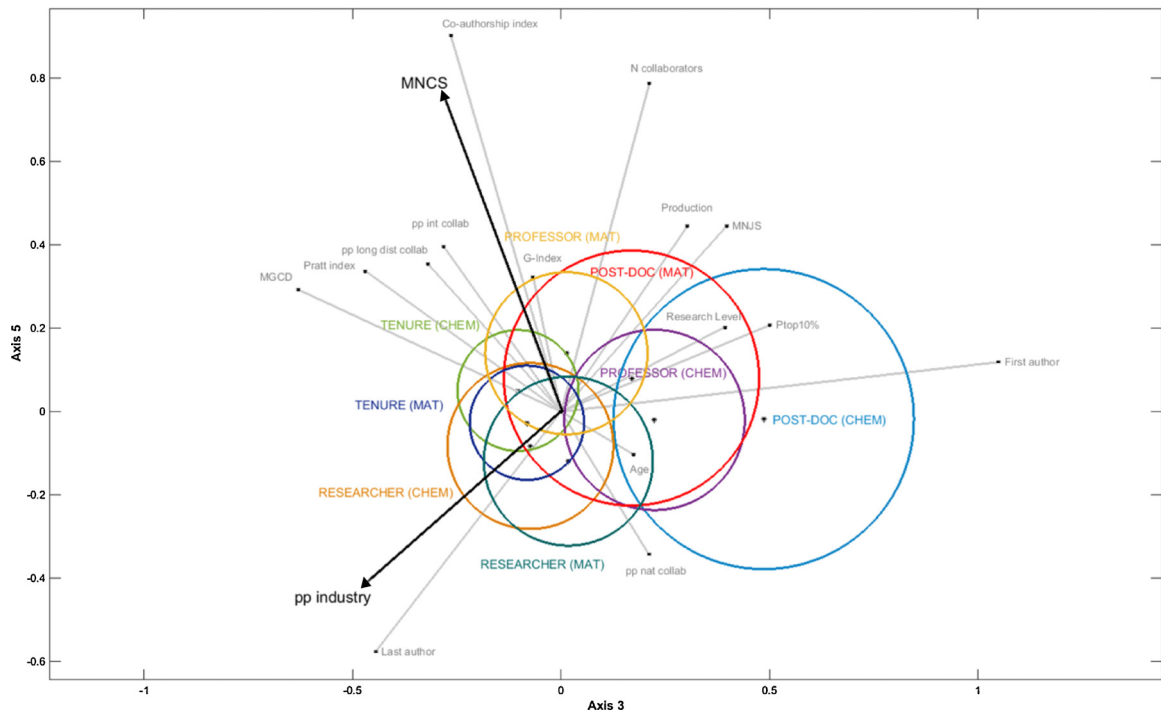


Fig. 3. Two dimensional Canonical Biplot for authors performance (plane 3–5). Variables coloured in grey present a QLR below 50 points.

power. Axis 3 is chiefly characterised by pp industry whereas Axis 5 accounts for MNCS. It is clear that all confidence intervals overlap, so there is no difference between groupings in terms of average citations and the degree of collaboration with enterprises. Besides, there is no relation between both indicators (right angle between vectors).

3.2. Effect sizes results

We have computed effect sizes by means of Hedge's g for some pairwise comparisons in order to estimate the magnitudes of the differences between groupings. These results are useful both due to the influence of sample size in the statistical values and as a means to support the Biplot analysis. Since the number of possible comparisons is extremely high, Table 4 displays a small but meaningful selection of effect sizes and variables. We mainly display effect sizes for Post-docs because of their smaller sample size.

Accordingly, the indicators which showed a stronger discriminatory power among groups of authors such as production, number of collaborators, first author and g -index obtain high g values as well. We can see what variables and pairs of groups differ most in terms of standard deviation units. In this regard, Post-doc researchers presented as a singular feature appear as first author in the byline at the Canonical Biplot. This becomes clear looking at Hedge's g for Post-docs with regard to the three types of full permanent scientists since size effects are larger ($g > 0.80$), which means the distribution of group scores

Table 4
Hedge's g by field and academic rank.

Academic rank	Production	N Collaborators	First author	g -index	pp int collab	Research level	MNCS	pp industry
Post-doc (CHEM) vs. Tenure (CHEM)	0.49	0.43	0.96	0.57	0.05	0.04	0.27	0.30
Post-doc (CHEM) vs. researcher (CHEM)	0.52	0.52	1.35	0.60	0.12	0.16	0.11	0.20
Post-doc (CHEM) vs. professor (CHEM)	0.52	0.53	1.41	0.58	0.06	0.18	0.14	0.20
Tenure (CHEM) vs. professor (MAT)	0.83	0.63	0.34	0.32	0.47	0.07	0.01	0.26
Researcher (CHEM) vs. tenure (MAT)	0.20	0.04	0.40	0.26	0.17	0.58	0.07	0.04
Post-doc (MAT) vs. tenure (MAT)	0.62	0.58	0.94	0.35	0.13	0.13	0.00	0.16
Post-doc (MAT) vs. researcher (MAT)	0.81	0.71	0.94	0.56	0.03	0.05	0.07	0.17
Post-doc (MAT) vs. professor (MAT)	0.91	0.83	1.60	0.57	0.23	0.17	0.07	0.01
Tenure (MAT) vs. professor (CHEM)	0.47	0.44	0.50	0.43	0.35	0.61	0.05	0.08
Researcher (MAT) vs. professor (CHEM)	0.29	0.30	0.36	0.23	0.27	0.53	0.01	0.06
Post-doc (MAT) vs. professor (CHEM)	0.50	0.56	1.22	0.53	0.22	0.47	0.06	0.12
Post-doc (CHEM) vs. professor (MAT)	0.94	0.78	1.83	0.64	0.53	0.12	0.16	0.10

* Effect size: $g < 0.20$ trivial, $g \geq 0.20$ small effect, $g \geq 0.50$ medium effect, $g \geq 0.80$ large effect and $g \geq 1.30$ very large effect.

do not overlap. Likewise, production, number of collaborators and *g*-index range from medium to large effect sizes which can be considered consistent enough. For instance, the largest effect sizes for production are observed between Post-docs and Research Professors, but also for Tenure (CHEM) vs. Professor (MAT) where $g = 0.83$. Interestingly, if we focus on international collaboration and research level, which were variables accounted for on Axis 2 of the Biplot, distributions within the same field are similar for all four academic ranks. Nevertheless, the biggest effect sizes arise when scientists from distinct fields are compared (Post-doc (CHEM) vs. Professor (MAT) $g = 0.53$ for pp int collab and Tenure (MAT) vs. Professor (CHEM) $g = 0.61$ for research level). On the other hand, MNCS and collaboration with industry, two of the indicators which showed weak discriminatory power for group means (no global differences) also present small *g* values for all pairwise comparisons, i.e. the differences are scant between groupings.

4. Discussion and conclusions

Due to the multidimensional nature of research activity, multiple indicators are needed to obtain a reliable picture of the research performance of scientists. Moreover, research performance may be affected by a wide range of factors, including personal data (age, gender), career (academic rank, promotion), organisational (teaching, mobility) or economic issues (research funds); and this generates a wide diversity of possibilities for the analysis of scientific performance. Due to this diversity and its multidimensionality, it is sometimes difficult to extract meaningful information from a multiplicity of variables and groups of researchers. In this study we provide a practical solution to this problem. Canonical Biplot analysis has been applied to study the behaviour of 729 researchers in two areas at the Spanish CSIC. Our aim was to build a joint plot for 8 groups (2 fields \times 4 academic ranks) and 17 variables where the groups were separated by the maximum discriminating power between them. As a result, we have studied which structure of variables offers the highest separation between groups of authors in a broad set of bibliometric indicators. Besides, statistical differences between group means have been assessed. Since test hypotheses are sensitive to different samples sizes, we have supported our results by means of effect sizes. They can be a relevant measure at the individual-level since data collection at this level is a tough job and it is not always possible to achieve a good sample size for the different categories or groupings.

In contrast to other multivariate techniques, Canonical Biplot offers some interesting and advantageous features. For instance, if we would have applied a MANOVA, we should have examined many tables and we would not have obtained a joint representation in a low dimensional space for a visual inspection of the underlying structure of the data matrix. If we would have used a Discriminant Analysis, we would have obtained a low dimensional plot describing the group's structure, but we would not have had direct information about the bibliometric indicators responsible for the separation between groupings and their correlations.

In our study, we have found two gradients for CSIC scientists. First, researchers of different academic ranks are clearly separated according to their age, level of production of papers, distinct number of collaborators, the number of highly-cited papers and their position in the byline. The second gradient relates to intrinsic field features since it separates Chemistry from Materials Science. The scientific activity of Materials Science researchers is less interdisciplinary and shows a higher share of international links compared to Chemistry, where the nature of research is more basic and takes place in larger teams. The lower interdisciplinarity of Materials Science is an unexpected finding, since it is a more recent and more applied field and it is supposed to require more diverse inputs. Actually, interdisciplinarity is often associated with application-oriented research and complex problem solving (van Rijnsvoever & Hessels, 2011). A possible explanation is that our interdisciplinarity measure considers the diversity of WoS categories to which the publication journals are assigned, but not the heterogeneous nature of each category. Accordingly, taking into account other measures such as the specialisation of team members could be advisable.

Our results show that the higher the academic rank, the higher the production of the researcher and his/her share of highly-cited papers and the more likely it is to appear as last author in the byline. The higher production of researchers in the highest academic rank has been described for other research areas at the Spanish CSIC (Costas et al., 2010) and elsewhere (Abramo, D'Angelo, & Di Costa, 2011b). Moreover, Abramo et al. (2014) denoted a moderate correlation for several fields (mild for Chemistry) between being a top productive scientist and the probability of having produced highly-cited papers. The same study also evidences that the highest probability of yielding highly-cited papers for non-top productive scientists lies with Assistant Professors, which can be considered to be on the same level as our Post-doc scientists. Interestingly, although our results show no differences in MNCS by academic rank and field, Post-docs obtain the highest average MNJS, probably because they are at the beginning of their careers and are very aware of the importance of publishing in prestigious journals to build a solid scientific career and gain a tenured position. On the other hand, an enhanced impact has been identified for those scientists who have heterogeneous collaboration patterns, chiefly at the international level (Franceschet & Costantini, 2010; Abramo et al., 2011a; Bordons, Aparicio, & Costas, 2013), even though there are some exceptions, maybe due to hyper-authorship (Cronin, 2001). Our results show that there are differences on international and long distance collaboration, but they seem to be more related to field adscription. However, having an extensive network of collaborators proved valuable in terms of performance.

As regards authorship order and age, it has been shown in the literature that, in experimental sciences, first-authored papers predominate among younger researchers while last position tends to be reserved to the more experienced scientists (Gingras, Larivière, Macaluso, & Robitaille, 2008; Costas & Bordons, 2011). Last-author position has been pointed out as a major position since it can be assumed to name the driving force behind the research, both intellectually and financially

(Tscharrntke et al., 2007). This is the role mainly played by authors in the higher academic ranks. Nevertheless, the strongest contribution to the actual work carried out is often expected from the researcher in the first-author position of the byline (Liu & Fang, 2014). This is indeed the most distinctive feature of Post-docs in our study, since they obtain the highest share of first-authored papers (around 1/3 of their production) and the lowest proportion of last-authored ones. Post-docs might be able to lead a research line and can work independently or within an existing team. Since we are dealing with two experimental fields, the junior signing pattern observed for Post-docs (Costas & Bordons, 2011) and their relatively high co-authorship index suggest their integration into more or less consolidated teams, which in fact is being fostered by some programmes to increase the size and competitiveness of teams. Although Post-docs are not so well connected to other people as researchers in tenured positions and yield less number of papers, they obtain an outstanding impact as measured by the MNJS. These results are consistent with previous studies focused on academic rank influence and signature order for CSIC scientists (see Costas & Bordons, 2011), where a relation between publishing in prestigious journals or being highly rewarded with citations and any specific position in the byline was not found.

We have to mention some limitations to the present study. First, the fact that each scientist was considered in the professional rank he/she held in 2007 and some of them may have been promoted during our period of reference. Nonetheless, this is only the case in a very small percentage of scientists. Second, due note should be taken that the small sample size of Post-docs limits the scope of the analysis and the significance of the results concerning this category of researchers. And third and last, our study focuses on CSIC scientists in two fields and we cannot generalise our results across other fields, institutions or countries. Concerning future research, including other factors such as promotion, teaching or mobility issues might be helpful to better understand certain patterns with a bearing on performance. PhD students also remain as an attractive group for further study due to the lack of large-scale analyses on them (Larivière, 2012, is an exception).

In summary, the approach set forth in this paper concerning the study of research performance at the individual level by means of a Canonical Biplot analysis enables us to examine a large set of indicators and to explore the underlying matrix structure. In this specific case, the most distinctive patterns that characterise researchers grouped by field and academic rank have been revealed. In view of all the foregoing, we conclude that the Canonical Biplot analysis is a strong exploratory tool with high potential in order to make headway in the unravelling of the intricate structure of relationships between research performance indicators and the individual characteristics of researchers.

Acknowledgements

This research was supported by the Spanish National Research Council (JAE pre-doctoral grant). This paper was conceived while Adrián A. Díaz-Faes carried out a research stay at the Centre for Science and Technology Studies (CWTS), Leiden University (The Netherlands).

References

- Abramo, G., D'Angelo, C. A., & Solazzi, M. (2011). Are researchers that collaborate more at the international level top performers? An investigation on the Italian university system. *Journal of Informetrics*, 5, 204–211. <http://dx.doi.org/10.1016/j.joi.2010.11.002>
- Abramo, G., D'Angelo, C. A., & Di Costa, F. (2011). Research productivity: Are higher academic ranks more productive than lower ones? *Scientometrics*, 88, 915–928.
- Abramo, G., Cicero, T., & D'Angelo, C. A. (2013). Individual research performance: A proposal for comparing apples to oranges. *Journal of Informetrics*, 7(2), 528–539. <http://dx.doi.org/10.1016/j.joi.2013.01.013>
- Abramo, G., Cicero, T., & D'Angelo, C. A. (2014). Are authors of highly cited articles also the most productives ones? *Journal of Informetrics*, 8, 89–97. <http://dx.doi.org/10.1016/j.joi.2013.10.011>
- Amaro, I. R., Vicente-Villardón, J. L., & Galindo, M. P. (2004). MANOVA Biplot para arreglos de tratamientos con dos factores basado en modelos lineales generales multivariantes. *Interciencia*, 29(1), 26–32.
- Bales, M. E., Dine, D. C., Merrill, J. A., Johnson, S. B., Bakken, S., & Weng, C. (2014). Associating co-authorship patterns with publications in high-impact journals. *Journal of Biomedical Informatics*, 52, 311–318.
- Bonacorsi, A. (2008). Search regimes and the industrial dynamics of science. *Minerva*, 46, 285–315.
- Bordons, M., Aparicio, J., & Costas, R. (2013). Heterogeneity of collaboration and its relationship with research impact in a biomedical field. *Scientometrics*, 96, 443–466.
- Bozeman, B., & Gaughan, M. (2011). How do men and women differ in research collaborations? An analysis of the collaborative motives and strategies of academic researchers. *Research Policy*, 40(1393), 1402. <http://dx.doi.org/10.1016/j.respol.2011.07.002>
- Cárdenas, O., Galindo, M. P., & Vicente-Villardón, J. L. (2007). Los métodos Biplot: evolución y aplicaciones. *Revista Venezolana de Análisis de Coyuntura*, 13(1), 279–303.
- Carpenter, M. P. (1979). Similarity of Pratt's measure of class concentration to the Gini index. *Journal of the American Society for Information Science*, 30(2), 108–110.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Costas, R., van Leeuwen, T. N., & Bordons, M. (2010). A bibliometric classificatory approach for the study and assessment of research performance at the individual level: The effects of age on productivity and impact. *Journal of the American Society for Information Science and Technology*, 61(8), 1564–1581. <http://dx.doi.org/10.1002/asi.21348>
- Costas, R., & Bordons, M. (2011). Do age and professional rank influence the order of authorship in scientific publications? Some evidence from a micro-level perspective. *Scientometrics*, 88, 145–161.
- Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, 52(7), 558–569.
- Díaz-Faes, A. A., Benito-García, N., Martín-Rodero, H., & Vicente-Villardón, J. L. (2011). Propuesta de aplicabilidad del método multivariante gráfico Biplot a los estudios bibliométricos en biomedicina. *Actas XIV Jornadas Bibliosalud*, 66. Retrieved from: (<http://hdl.handle.net/10760/15998>) (Cádiz, España: BV-SSPA).
- Díaz-Faes, A. A., González-Albo, B., Galindo, M. P., & Bordons, M. (2013). HJ-Biplot como herramienta de inspección de matrices de datos bibliométricos. *Revista Española de Documentación Científica*, 36(1), e001. <http://dx.doi.org/10.3989/redc.2013.1.988>

- Del Re, A. C. (2013). *Compute.es: Compute effect sizes. R package version 0.2-2*. Retrieved from: (<http://cran.r-project.org/web/packages/compute.es>).
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152.
- Franceschet, M., & Costantini, A. (2010). The effect of scholar collaboration on impact and quality of academic papers. *Journal of Informetrics*, 4, 540–553.
- Gabriel, K. R. (1971). The Biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58(3), 453–467.
- Gabriel, K. R. (1972). Analysis of meteorological data by means of canonical decomposition and Biplots. *Journal of Applied Meteorology*, 11, 1071–1077.
- Gabriel, K. R., & Odoroff, C. L. (1990). Biplots in biomedical research. *Statistics in Medicine*, 9, 469–485. <http://dx.doi.org/10.1002/sim.4780090502>.
- Galindo, M. P. (1986). Una alternativa de representación simultánea: HJ-Biplot. *Questió*, 10(1), 13–23.
- Gingras, Y., Larivière, V., Macaluso, B., & Robitaille, J. P. (2008). The effects of aging on researchers' publication and citation patterns. *PLoS ONE*, 3(12), e4048.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 106–128.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–16572.
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56(5), 746–759.
- Larivière, V. (2012). On the shoulders of students? The contribution of PhD students to the advancement of knowledge. *Scientometrics*, 90(2), 463–481. <http://dx.doi.org/10.1007/s11192-011-0495-6>
- Liu, X. Z., & Fang, H. (2014). Scientific group leaders' authorship preferences: An empirical investigation. *Scientometrics*, 98, 909–925.
- Martin, B. R. (1996). The use of multiple indicators in the assessment of basic research. *Scientometrics*, 36(3), 343–362.
- Moravcsik, M. J. (1984). Life in a multidimensional world. *Scientometrics*, 6(2), 75–86.
- Morillo, F., Díaz-Faes, A. A., González-Albo, B., & Moreno, L. (2014). Do networking centres perform better? An exploratory analysis in psychiatry and gastroenterology/hepatology in Spain. *Scientometrics*, 98(2), 1401–1416. <http://dx.doi.org/10.1007/s11192-013-1183-5>
- Narin, F., Pinski, G., & Gee, H. H. (1976). Structure of the biomedical literature. *JASIS*, 27(1), 25–44.
- Pratt, A. D. (1977). A measure of class concentration in bibliometrics. *JASIS*, 28(5), 285–292.
- Radicchi, F., & Castellano, C. (2013). Analysis of bibliometric indicators for individual scholars in a large data set. *Scientometrics*, 97(3), 627–637. <http://dx.doi.org/10.1007/s11192-013-1027-3>
- Ruiz-Castillo, J., & Costas, R. (2014). The skewness of scientific productivity. *Journal of Informetrics*, 8, 917–934.
- Schneider, J. W. (2013). Caveats for using statistical significance tests in research assessments. *Journal of Informetrics*, 7(1), 50–62. <http://dx.doi.org/10.1016/j.joi.2012.08.005>
- Seiler, C., & Wohlrabe, K. (2013). Archetypal scientists. *Journal of Informetrics*, 7(2), 345–356. <http://dx.doi.org/10.1016/j.joi.2012.11.013>
- Sonnenwald, D. H. (2007). Scientific collaboration. *Annual Review of Information Science and Technology*, 41, 643–681.
- Tijssen, R. J. W. (2011). Joint research publications: A performance indicator of university-industry collaboration. *Evaluation in Higher Education*, 5(2), 19–40.
- Torres-Salinas, D., Robinson-García, N., Jiménez-Contreras, E., Herrera, F., & Delgado López-Cózar, E. (2013). On the use of biplot analysis for multivariate bibliometric and scientific indicators. *Journal of the American Society for Information Science and Technology*, 64(7), 1468–1479.
- Tscharntke, T., Hochberg, M. E., Rand, T. A., Resh, V. H., & Kruass, J. (2007). Author sequence and credit for contributions in multi-authored publications. *PLoS Biology*, 5(1), 13–14.
- van Rijnsoever, F. J., & Hessels, L. K. (2011). Factors associated with disciplinary and interdisciplinary research collaboration. *Research Policy*, 40, 463–472.
- Varas, M. L., Vicente-Tavera, S., Molina, E., & Vicente-Villardón. (2005). Role of canonical Biplot method in the study of building stones: An example from Spanish monumental heritage. *Environmetrics*, 16, 1–15.
- Vicente-Villardón, J. L. (1992). *Una alternativa a los métodos factoriales clásicos basada en una generalización de los métodos biplot*. Salamanca, Spain: University of Salamanca (PhD thesis).
- Vicente-Villardón, J. L. (2014). *MultBiplot: A package for Multivariate Analysis using Biplots*. Departamento de Estadística, Universidad de Salamanca. Retrieved from: (<http://biplot.usa.es/ClassicalBiplot/index.html>) [Software].
- Waltman, L., Tijssen, R. J. W., & van Eck, N. J. (2011). Globalisation of science in kilometres. *Journal of Informetrics*, 5(4), 574–582. <http://dx.doi.org/10.1016/j.joi.2011.05.003>
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37–47. <http://dx.doi.org/10.1007/s11192-011-0354-5>
- Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, 6(4), 700–711. <http://dx.doi.org/10.1016/j.joi.2012.07.008>