

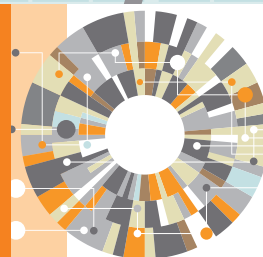


Elsevier Research Intelligence

SciVal

SciVal Metrics Guidebook

Version 1.01 | February 2014



Empowering Knowledge

Authored by

Dr Lisa Colledge and Dr Reinder Verlinde

**Version 1.01
February 2014**

Table of contents

1.0	Scopus: the data source for SciVal metrics	7	3.2.1	Size	27
1.1	Scopus content	7	3.2.2	Discipline	28
1.1.1	Scopus content and SciVal	7	3.2.3	Publication-type	31
1.1.2	Books and SciVal	7	3.2.4	Database coverage	32
1.2	Data currency in Scopus and SciVal	8	3.2.5	Manipulation	34
1.3	Author Profiles	8	3.2.6	Time	34
1.3.1	Author Profiles in Scopus	8	4.0	SciVal metrics: methods and use	35
1.3.2	Authors and Researchers in SciVal	9	4.1	The display of metrics in SciVal	38
1.4	Affiliation Profiles	10	4.2	Metric: Scholarly Output	39
1.4.1	Affiliation Profiles in Scopus	10	4.3	Metric: Journal Count	41
1.4.2	Affiliations and Institutions in SciVal	10	4.4	Metric: Journal Category Count	43
1.5	Do publications belong to Institutions or to Researchers?	11	Example 2:	Scholarly Output, Journal Count and Category Count	46
1.6	Organization-types	11	4.5	Metric: Citation Count	48
1.6.1	Organization-types in Scopus	11	4.6	Metric: Cited Publications	51
1.6.2	Organization-types in SciVal	11	4.7	Metric: Citations per Publication	53
1.7	Journal metrics in Scopus and SciVal	12	Example 3:	Citation Count, Cited Publications and Citations per Publication	56
2.0	SciVal and Metrics	13	4.8	Metric: Number of Citing Countries	58
2.1	Groups of Metrics in SciVal	13	Example 4:	Number of Citing Countries	60
2.2	The calculation and display of metrics in SciVal	14	4.9	Metric: Field-Weighted Citation Impact	61
2.2.1	Publications included in the calculation of a metric	14	Example 5:	Field-Weighted Citation Impact	64
2.2.2	Deduplication	14	4.10	Metric: Collaboration	66
2.2.3	Zero and null values	15	4.11	Metric: Collaboration Impact	69
2.2.4	The display of “>current year”	15	4.12	Metric: Academic-Corporate Collaboration	71
2.2.5	Citation counts	16	4.13	Metric: Academic-Corporate Collaboration Impact	73
2.2.6	Calculation options	17	Example 6:	Collaboration, Collaboration Impact, Academic-Corporate Collaboration, and Academic-Corporate Collaboration Impact	76
2.2.6.1	Research Area filter	18	4.14	Metric: Outputs in Top Percentiles	78
2.2.6.2	Publication-type filter	18	Example 7:	Outputs In Top Percentiles	82
2.2.6.3	Self-citation exclusion	19	4.15	Metric: Publications in Top Journal Percentiles	84
2.2.6.4	Absolute number and percentage options	19	Example 8:	Publication In Top Journal Percentiles	88
Example 1a:	Self-Citation Exclusion	20	4.16	Metric: <i>h</i> -indices	90
Example 1b:	Self-Citation Exclusion	22	Example 9:	<i>h</i> -Indices	94
3.0	Selection of appropriate metrics	25			
3.1	Clarity on the question being asked	26			
3.2	Factors besides performance that affect the value of a metric	26			

Foreword

The quote often attributed to Albert Einstein, but perhaps more properly attributed to William Bruce Cameron¹, is often referred to when writing a foreword such as this: “Not everything that counts can be counted, and not everything that can be counted counts”.

This is undoubtedly true, but does not follow that nothing should be measured.

There is much that can be counted that is important and provides valuable perspectives on trends in academia, and there is an increasing emphasis on this in the world of research today.

The specialized work of engineers is used every day by car owners, who know how to use the technology that has been developed by others in an appropriate and responsible way. Car owners welcome the continued contributions of engineers to their vehicles, since it makes their lives easier. In the same way, the field of metrics relies on specialized academic study for its development and to capitalize on advances in technology; these scholarly outputs are increasingly being used by people involved in research, whether directly by conducting research or indirectly in a supporting or enabling role. The majority of these people would not consider themselves to be experts in metrics or their use: this Guidebook is aimed at those people, and at supporting the use of metrics in an appropriate and responsible way.

This Guidebook is not an academic tome that aims to provide up-to-the-minute information about the latest advances in metrics. It is intended to be a straightforward, practical companion to the use of SciVal, which is part of the Elsevier Research Intelligence suite; new versions of this Guidebook will be produced as SciVal continues to evolve. It provides some facts about how the data underlying the metrics are used, about how the metrics are calculated and displayed, and about variables besides performance that can affect the values of these metrics. It then makes some suggestions about situations in which the metrics are useful, when care should be taken, and how shortcomings may be addressed.

Only one absolute in the use of metrics is highlighted in this Guidebook: always use more than 1 metric to give insights into a question, and always support conclusions drawn from metrics by peer review and/or expert opinion. This “triangulation” of approaches will increase the reliability of conclusions drawn if these distinct types of information reinforce each other, and will flag areas for further investigation if they do not. Beyond this, there are no black-and-white rules. The best approach is to use common sense.

I hope that you find this Guidebook useful, and perhaps even interesting.

Dr Lisa Colledge
Elsevier
January 2014

1: <http://quoteinvestigator.com/2010/05/26/everything-counts-einstein/>

1. Scopus: the data source for SciVal metrics

All metrics and information displayed in SciVal at the time of writing this Guidebook are based on Scopus. This section highlights some of the approaches of Scopus that are useful in understanding the metrics in SciVal. The reader can find more extensive information about Scopus online².

1.1 Scopus content

The independent and international Scopus Content Selection and Advisory Board reviews titles for inclusion in Scopus on a continuous basis. Information about the process and the acceptance criteria is available online³. The Board considers journals, conference proceedings, trade publications, book series, and stand-alone books for inclusion.

Scopus indexes about 50 million publications. Reference lists are captured for the 29 million records published from 1996 onwards. The additional 21 million pre-1996 records reach as far back as the publication year 1823.

1.1.1 Scopus content and SciVal

SciVal uses Scopus content from 1996 onwards so that the citation counts displayed in SciVal are based on uninterrupted years of data.

1.1.2 Books and SciVal

Scopus indexes both book series, and stand-alone books. “Books” in SciVal refers to stand-alone books only; their characteristics, and the fact that they do not have journal metrics, sets them apart from the series of books, journals, conference proceedings and trade publications.

Scopus links citations to the individual chapters of edited volumes when the information provided by the author allows this, and otherwise to the edited volume itself. Scopus makes either one link or the other, but not both. SciVal credits a Researcher who is the author of a book chapter with the count of citations linked to that particular chapter; it credits a Researcher who is the editor of the entire volume with the citations linked to the volume plus those linked to all the individual chapters to ensure that the editor is credited with the full citation impact of their scholarly contribution.

2: <http://www.elsevier.com/online-tools/scopus>

3: <http://www.elsevier.com/online-tools/scopus/content-overview>

1.2 Data currency in Scopus and SciVal

The data in Scopus are updated daily.

The data in SciVal are updated every week. SciVal takes a copy of the Scopus database that is then structured to optimally support its metrics and functionality. This means that SciVal may be slightly behind Scopus in its data currency.

1.3 Author Profiles

1.3.1 Author Profiles in Scopus

Scopus is the only database in the world which has invested in automatically grouping the publications it indexes into those published by a single author. Author Identifiers (Author Profiles) group together publications belonging to one author, and they have 2 modes of input:

- **Publications are automatically grouped into Author Profiles using a matching algorithm:**
 - This algorithm looks for similarities in author name, as well as affiliation, journal portfolio, and discipline to match publications together. Users may notice that multiple name variants are grouped within one Author Profile, which indicates the value of this algorithm
 - The information provided by authors is not always consistent or complete, and even if this were the case, the mobility of authors means that there is always some doubt about whether some publications belong together. In situations like these, a balance needs to be made between the precision, or accuracy, of matching, and the recall, or completeness, of the groups formed, and increasing one will reduce the other
 - The Scopus algorithm favors accuracy, and only groups together publications when the confidence level that they belong together, the precision of matching, is at least 99%, such that in a group of 100 papers, 99 will be correctly assigned. This level of accuracy results in a recall of 95% across the database: if an author has published 100 papers, on average 95 of them will be grouped together by Scopus
 - These precision and recall figures are accurate across the entire Scopus database. There are situations where the concentration of similar names increases the fragmentation of publications between Author Profiles, such as in the well-known example of Chinese authors. Equally there are instances where a high level of distinction in names results in a lower level of fragmentation, such as in Western countries
 - A publication that has multiple co-authors will be part of multiple Author Profiles
- **Publications are manually reassigned based on feedback.** The matching algorithm can never be 100% correct because the data it is using to make the assignments are not 100% complete or consistent. The algorithm is therefore supplemented by feedback received from the sector, including from the Open Researcher and Contributor ID (ORCID) initiative⁴, and that feedback is used to enhance the profiling of authors by the Scopus Author Feedback Team⁵

4: <http://orcid.org/>

5: <http://www.scopusfeedback.com/>

1.3.2 Authors and Researchers in SciVal

The presence of Author Profiles in Scopus enables enormous flexibility for SciVal users. Every Author Profile is available to SciVal users to define and view any Researcher in the world in real time, whether they are based at the user's institution, they are a collaboration partner in another country, or they are a Researcher who was unknown until today. Users can build on the Author Profiles to define as many Researchers as they like in their SciVal accounts, and subsequently use their Researchers as the basis to create an unlimited number of:

- **Groups of Researchers.** These could represent actual research teams, departments or other organizational groups, international collaboration networks, or models of research teams that are being considered
- **Publication Sets.** These are pieces of a Researcher's output, such as publications from a particular affiliation, or those funded by a particular award
- **Groups of Publication Sets.** These could represent a selection of publications being submitted to a national evaluation, such as the Research Excellence Framework in the United Kingdom⁶, for example

SciVal distinguishes between its use of the terms “authors” and “Researchers”:

- **Authors are the automatically created Scopus Author Profiles.** The author count in the Overview module, for example, is a count of unique Author Profiles; this may be an over-estimation of the number of Researchers, because the recall rate of 95% means that a Researcher's publication portfolio might be split over multiple Author Profiles
- **Researchers are entities which have been created with the benefit of human input.** SciVal users can combine Author Profiles, remove any publications that do not belong to an Author Profile, and search for individual publications that should be added. The Profile Refinement Service that populates SciVal with Researchers and Groups of Researchers on behalf of an Institution is a similar manual process, performed by Elsevier on behalf of a customer. SciVal distinguishes these manually created Researchers, whether created by users or by Elsevier, from the automatically generated Author Profiles. All of these manual enhancements are fed back to Scopus and used to improve the quality of the source database for all users of Scopus data. This means that once the feedback has been processed by Scopus, each Researcher in SciVal is represented by a single Author Profile which can be automatically updated as new publications are indexed

6: <http://www.ref.ac.uk/>

1.4 Affiliation Profiles

1.4.1 Affiliation Profiles in Scopus

Scopus is the only database in the world which has invested in automatically grouping the publications it indexes into those published by a single affiliation. These groups of publications belonging to one affiliation are called Affiliation Profiles, and they have 2 modes of input:

- **Publications are automatically grouped into Affiliation Profiles using a matching algorithm:**
 - This algorithm looks for similarities in affiliation name, as well as addresses, to match publications together. Users may notice that multiple name variants are grouped within one Affiliation Profile, which indicates the value of this algorithm. Scopus makes use of an authoritative database that contains over 70,000 manually verified institutional name variants to match publications together
 - The information provided by authors is not always consistent or complete, so that there is always some doubt about whether some publications belong together; in situations like these, a balance needs to be made between the precision, or accuracy, of matching, and the recall, or completeness, of the groups formed, and increasing one will reduce the other
 - The Scopus algorithm favors accuracy, and only groups together publications when the confidence level that they belong together, the precision of matching, is at least 99%, such that in a group of 100 papers, 99 will be correctly assigned. This results in a recall of 93% across the database, such that if an Affiliation has published 100 papers, on average 93 of them will be grouped together by Scopus, and the others will be in 1 or more separate groups
 - A publication that has co-authors with multiple affiliations will be part of multiple Affiliation Profiles
- **Publications are manually reassigned based on feedback.** The matching algorithm can never be 100% correct because the data it is using to make the assignments are not 100% complete or consistent. The algorithm is therefore supplemented by feedback received from the official authority of the affiliation in question

1.4.2 Affiliations and Institutions in SciVal

The presence of Affiliation Profiles in Scopus brings enormous benefits for SciVal users, with metrics being pre-calculated and available to view at the click of a mouse. SciVal users also benefit from the availability of **Groups of Institutions**, such as American states.

SciVal distinguishes between its use of the terms “affiliations” and “Institutions”:

- **Affiliations are the automatically created Scopus Affiliation Profiles.** A medical school will be a separate Affiliation Profile from a university, for instance
- **Institutions are groupings of related Affiliation Profiles** which have been manually created as a convenient starting point for SciVal users; approximately 4,500 Institutions have been pre-defined and are available in SciVal. Any medical schools are always grouped together with the university in the SciVal Institutions

1.5 Do publications belong to Institutions or to Researchers?

Researchers are mobile, and tend to change affiliations during their careers. This leads to 2 perspectives as to whether publications “belong” to Institutions or to Researchers:

- The “**Institutional Perspective**” is typically that publications should remain assigned to them even when the Researchers that authored these publications have moved. In other words, the publications are not mobile, despite their authors moving around
- The “**Researcher Perspective**” is generally that publications should be just as mobile as their authors, and should move from affiliation to affiliation as their authors’ careers develop

These approaches are both needed to completely support questions that are asked in different situations. SciVal offers both of these perspectives, because publications are linked to both Affiliation and Author Profiles independently of each other:

- **Institutions and Groups of Institutions in SciVal** take the “Institutional Perspective”
- **Researchers, Groups of Researchers, Publications Sets and Groups of Publication Sets in SciVal** take the “Researcher Perspective”

1.6 Organization-types

1.6.1 Organization-types in Scopus

Organization-types are assigned to Scopus Affiliation Profiles based on their primary functions. This function is often very clear from the name of the affiliation, and the organization’s website is checked for guidance if there is any doubt.

Scopus assigns affiliations to the following organization-types: university, college, medical school, hospital, research institute, corporate, law firm, government, military organization, and non-governmental organization.

1.6.2 Organization-types in SciVal

The organization-types used in SciVal are based on aggregations of the Scopus organization-types to group similar functions together, and to simplify the options for the user. SciVal uses 5 organization-types: Academic, Corporate, Government, Medical, and Other. These are composed of the following Scopus organization-types:

- **Academic:** university, college, medical school, and research institute
- **Corporate:** corporate and law firm
- **Government:** government and military organization
- **Medical:** hospital
- **Other:** non-governmental organization

1.7 Journal metrics in Scopus and SciVal

Scopus and SciVal use 2 journal metrics that have been developed by academic research teams, and whose methodology has been published in peer-reviewed journals. Information about these metrics is available online, and all the metrics values are also available for free⁷. These metrics are:

- **SNIP⁸. Source-Normalized Impact per Paper** is a ratio between the “Raw Impact per Paper”, a type of Citations per Publication calculation, actually received by the journal, compared to the “Citation Potential”, or expected Citations per Publication, of that journal’s field. SNIP takes differences in disciplinary characteristics into account, and can be used to compare journals in different fields. The average SNIP value for all journals in Scopus is 1.000
- **SJR⁹. SCImago Journal Rank** is a prestige metric, whose methodology is similar to that of Google PageRank. It weights the value of a citation depending on the field, quality and reputation of the journal that the citation comes from, so that “all citations are not equal”. SJR also takes differences in the behavior of academics in different disciplines into account, and can be used to compare journals in different fields. The average SJR value for all journals in Scopus is 1.000

Journal metrics are not calculated for trade publications. Stand-alone books cannot have SNIP and SJR values, since their calculation methodologies depend on items being in a series.

7: www.journalmetrics.com

8: <http://www.sciencedirect.com/science/article/pii/S1751157710000039>

9: <http://www.sciencedirect.com/science/article/pii/S1751157710000246>

2. SciVal and Metrics

2.1 Groups of Metrics in SciVal

SciVal offers a broad range of metrics to:

- Accommodate the preferences of users in approaching questions from multiple angles
- Enable users to “triangulate” their evidence. When seeking input into a question from an evidence-base, it is preferable to use at least 2 different metrics to provide information; if pieces of intelligence gained from multiple metrics reinforce each other, then the user can have a higher degree of confidence that their conclusions are valid

SciVal’s metrics can be classified within 6 groups, and a metric may be part of more than 1 group, as illustrated in **Table 1**:

- **Productivity metrics** give information on the volume of output of an entity
- **Citation Impact metrics** indicate the influence of an entity’s output, as indicated by various types of citation counts
- **Collaboration metrics** provide information on the research partnerships of an entity
- **Disciplinarity metrics** give information on the spread of topics within an entity’s publications
- **Snowball Metrics**¹⁰ are defined and endorsed by research-intensive universities as providing important insight into institutional strategies:
 - The agreed and tested definitions are shared free-of-charge with the sector in the Snowball Metrics Recipe Book¹¹, with the ambition that Snowball Metrics become global standards for the higher education sector
 - These “recipes” can be used by anyone for their own purposes. Elsevier supports Snowball Metrics as recognized industry standards, and is implementing these metrics in relevant systems and tools, including SciVal
 - Snowball Metrics are indicated in the SciVal interface by the following snowflake symbol ❄️
- **“Power metrics”** whose value tends to increase as the size of an entity increases. For example, a larger institution will tend to publish more output than a smaller institution, just because of its greater size

10: www.snowballmetrics.com

11: www.snowballmetrics.com/metrics

	Productivity	Citation Impact	Collaboration	Disciplinarity	Snowball Metric	“Power metric”
Scholarly Output						
Journal Count						
Journal Category Count						
Citation Count						
Cited Publications						
Citations per Publication						
Number of Citing Countries						
Field-Weighted Citation Impact						
Collaboration						
Collaboration Impact						
Academic-Corporate Collaboration						
Academic-Corporate Collaboration Impact						
Outputs in Top Percentiles						
Publications in Top Journal Percentiles						
<i>h</i> -indices						

Table 1: Groups of metrics in SciVal. Those metrics that have half a shaded cell in the “Power metric” column are “Power metrics” when the “Absolute number” option is selected, but not when the “Percentage” option is selected.

2.2 The calculation and display of metrics in SciVal

2.2.1 Publications included in the calculation of a metric

The ideal situation would be that every publication in a data set is associated with the information needed so that it can be included in the calculation of every metric. In practice this is not the case; authors do not always include complete affiliation information, and publications are not always part of items indexed in Scopus that have journal metrics values, for example. **Publications that lack the necessary information are excluded from the metric calculation.**

2.2.2 Deduplication

SciVal offers the user the opportunity to investigate aggregate entities formed by the combination of smaller entities; for example, US states are groups of institutions, and geographical regions are groups of countries.

The same publication may be part of multiple smaller component entities, and can be added multiple times to the aggregate entity. Say, for example, that the Researcher R1 has co-authored a publication P1 with Researcher R2; in that case, P1 is a part of both R1 and R2.

SciVal deduplicates all the publications within an aggregate entity, so that a publication is only counted once even if it is co-authored by several of the component entities. **Entities are groups of unique publications**, and users can have confidence when creating aggregate entities that SciVal deduplicates the data set. In the above example, P1 is counted once only in a Group of Researchers that is composed of R1 and R2.

2.2.3 Zero and null values

SciVal distinguishes between zero and null values (absence of a value), which have distinct meanings for the interpretation of metrics.

Consider the metric Scholarly Output, which calculates the number of publications of an entity. If an entity has not published any outputs in a particular time frame, this is displayed as a zero value and the lack of publication is important to understanding an entity's activity. Null values are never displayed for Scholarly Output.

This is not the case for a metric like Citation Count, which counts the number of citations that an entity's publications have received. It is not meaningful to display a value of zero citations for a period during which an entity has not published anything, because the entity simply cannot receive citations when there is an absence of publications. This is different from the situation when an entity has published items that have not received any citations.

SciVal therefore distinguishes between the following 2 cases:

- If an entity has published output in a particular period, and that output has not received any citations, then a value of zero is displayed for Citation Count
- If an entity has not published any output in a particular time frame, then the entity cannot receive any citations during that period. In this case, a null value is plotted in SciVal; users may notice a break in a line when viewing metrics over time in the Benchmarking module, for example

This same reasoning is applied to other metrics. It is not possible to be internationally collaborative, for example, if nothing has been published.

Outputs in Top Percentiles can only be calculated for the current year from the first data snapshot on or after 1 July. It will be displayed as a null value until this date is reached. This metric depends on being able to divide the publications into 100 percentiles, and this level of division is not possible earlier in the publication year when items just published have received very few citations.

2.2.4 The display of ">current year"

A high proportion of the items indexed in Scopus are journals. Publishers sometimes release journal issues before their official cover date; for instance, at the end of 2013, users can find items in the Scopus database that belong to the 2014 publication year, and occasionally even to later years.

A great deal of importance is placed in SciVal on a high currency of data, so that all information present in the database is available to users. All publications that belong to items with a cover date later than the current year are collected together in the "Publication Year" labelled ">2013" in 2013, ">2014" in 2014, and so on.

	Research Area filter	Publication-type filter	Self-citation exclusion	Absolute count or percentage
SciVal Default	No filter applied	No filter applied	No exclusion	Percentage
Scholarly Output				
Journal Count				
Journal Category Count				
Citation Count				
Cited Publications				
Citations per Publication				
Number of Citing Countries				
Field-Weighted Citation Impact				
Collaboration				
Collaboration Impact				
Academic-Corporate Collaboration				
Academic-Corporate Collaboration Impact				
Outputs in Top Percentiles				
Publications in Top Journals Percentiles				
<i>h</i> -indices				

Table 2: Calculation options for the metrics in SciVal

2.2.5 Citation counts

Citations counts in SciVal represent the total number of citations received since an item was published, up to the date of the last data cut. Citation Impact metrics are often displayed in a chart or table with years in SciVal to indicate trends; these years are always the years in which items were published, and never refer to the years in which citations were received.

Older publications tend to have more citations than newer publications, simply because they have had longer to receive them from subsequent work. This means that for metrics like Citation Count and Citations per Publication, where this time factor is not accounted for, the user will notice a “dip” in the timeline in recent years. Comparisons between similar entities will still be useful and valid despite this dip, since the time factor will affect all similar entities equally, but if the user prefers to avoid this display they can select a metric like Field-Weighted Citation Impact which inherently accounts for this.

[illegible]

2.2.6 Calculation options

SciVal users have different needs in how they use the metrics. Some of these needs are met by how the results are visualized, but some affect the method of calculation. For instance:

- An academic who is a journal editor may well publish regular editorials, which count towards their productivity and may also increase their citation impact. In understanding the citation impact of this academic, it is very important to look at the entire body of their output, including the editorials, since being a journal editor and writing editorials is a significant part of their scholarly contribution. However, when comparing this academic to others who are not journal editors, it may sometimes be desirable to exclude the effect of the editorials to ensure a fair comparison
- Self-citations are those by which an entity refers to its previous work in new publications. Self-citing is normal and expected academic behavior, and it is an author's responsibility to make sure their readers are aware of related, relevant work. However, some users prefer to exclude self-citations from Citation Impact metrics, whereas to others this is not critical to the analysis being done

There is no universally right or wrong approach to these calculation preferences, although in some situations an option may become very important. These calculation preferences are available in SciVal as options to the user, some of which are described below. Each metric has its own set of calculation options, which are summarized in [Table 2](#).

2.2.6.1 Research Area filter

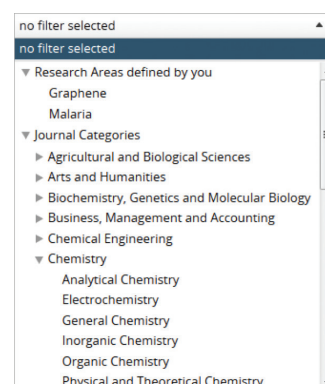
The Research Area filter limits the metric calculations to the publications of an entity that fall within a particular discipline. Where journal classifications are used, as in the accompanying screenshot, both the Main Categories and Sub-Categories are available to use as Research Area filters. User-defined Research Areas can also be used in some modules, such as the Research Areas “Graphene” and “Malaria” in the screenshot.

The Research Area filter applies to the entity’s publications only. It does not have any effect on the citation counts used to calculate the Citation Impact metrics; the citations are counted regardless of the Research Area of the citing paper.

Some metrics are not field-normalized. This means that differences in the behavior of academics in distinct disciplines that can affect metrics values are not accounted for in their calculation, and it may be difficult for a user to distinguish between differences in disciplinary behavior and true differences in research activity. These non-field-normalized metrics are very useful when comparing entities that fall into the same discipline, but it is not advisable to use non-field-normalized metrics to compare entities that fall into different fields. Users can address this by:

- **Using the Research Area filter** when using a non-field-normalized metric to compare entities made up of a mixture of disciplines, such as an Institution or a Country. This has the effect of enabling benchmarking of comparable disciplinary slices of these entities
- **Using field-normalized metrics** such as Field-Weighted Citation Impact which inherently take disciplinary differences into account

The default setting in SciVal is that no Research Area filter is applied.



2.2.6.2 Publication-type filter

The Publication-type filter limits the metric calculations to articles, reviews, and/or books for example. This filter can be applied when the user judges that this variable is important for their analysis, such as:

- Distinguishing between original research contributions that are often published as articles, and expert opinion, typically communicated in reviews
- In disciplines such as Engineering and Computer Science, where it is sometimes important to focus on conference proceedings
- When comparing an academic who is a journal editor and has the opportunity to publish editorials, with an academic who is not a journal editor

The Publication-type filter applies to the entity’s publications only. It does not have any effect on the citation counts used to calculate the Citation Impact metrics; the citations are counted regardless of the Publication-type of the citing paper. The only exception is Outputs in Top Percentiles, where this filter also applies to the data universe used to generate the citation thresholds.

The default setting in SciVal is that no Publication-type filter is applied.

2.2.6.3 Self-citation exclusion

Self-citations are those by which an entity refers to its previous work in new publications. Self-citations are typically thought of in the context of an individual academic, but journal self-citations, institutional self-citations, and country-self-citations are examples of the same activity at different levels.

There is nothing inherently wrong with the act of self-citing. It is normal and expected academic behavior to build on work previously conducted and published, and to refer in new publications to older material to alert the reader to important and relevant information that will aid their understanding. Indeed, the most relevant work to be aware of may well be research that has previously been published by the same entity, and it is an author's responsibility to make their readers aware of it.

The act of referring to previous work is, however, open to abuse, and there are rare cases of unethical self-citation practices which have led to self-citation acquiring a somewhat bad reputation. SciVal users have the option to exclude self-citations from the calculation of many of the Citation Impact metrics, to make their own judgment about whether self-citation rates are within normal boundaries. SciVal distinguishes between self-citations at the author, institutional, and country levels, and applies the appropriate exclusion depending on the type of entity being viewed without the user needing to specify this.

The default setting in SciVal is that all citations are included.

See [Examples 1a](#) and [1b](#) for illustrations of how self-citation exclusion is performed in SciVal.

2.2.6.4 Absolute number and percentage options

Some metrics in SciVal offer the option of being viewed as "Absolute number", or as "Percentage":

- Users are advised to select the **"Percentage" option to size-normalize**, when they are comparing entities of different sizes
- It is useful to select the **"Absolute number" option when it is important to communicate the magnitude** of the number of publications involved

The default setting in SciVal is "Percentage".

Example 1a: Self-Citation Exclusion

Scenario: The user is looking at a Country, Institution or Researcher entity that consists of 1 publication, PP. Say that this entity has received 6 citations from P1, P2, P4, P5, P6 and P7.

Publication	PP
Author	AA
Institution	II
Country	CC



PP is cited by

Publication	Author	Institution	Country
P1	A1	I1	C1
P2	AA	I2	C2
P4	AA	I3	CC
P5	AA	II	CC
P6	A2	II	CC
P7	A2	I3	CC

Question: What happens when the user chooses not to include self-citations when calculating a metric that offers this as an option?

Answer: If the entity you are viewing is Country CC, then the citations from P4, P5, P6 and P7 are self-citations because the affiliation country CC is the same as the entity's Country CC.

These 4 citations would no longer be included in the metric calculation.

Answer: If the entity you are viewing is Institution II, then the citations from P5 and P6 are self-citations because the affiliation institution II is the same as the entity's Institution II.

These 2 citations would no longer be included in the metric calculation.

Answer: If the entity you are viewing is a Researcher that includes author AA, then the citations from P2, P4 and P5 are self-citations because the author AA is the same as that of the entity's author AA.

These 3 citations would no longer be included in the metric calculation.

The citation from P1 is never classified as a self-citation, regardless of the type of entity.

Example 1b: Self-Citation Exclusion

Scenario: The user is looking at a Country, Institution or Researcher that consists of 3 publications, P4, P5 and P6:

- The publication P4 has 2 authors: A2 and A5. They are both based at institution I2, in country C2.
- The publication P5 has 2 authors: A1 and A3. Author A1 has joint affiliations: I1 and I3 in countries C1 and C2 respectively. A3 has a single affiliation: I2 in country C2.
- The publication P6 has the same 2 authors as P5: A1 and A3. Author A1 has again listed joint affiliations: A1 has remained at I3 in country C2, but the second affiliation is I5 in C4. A3 still has a single affiliation but it is different from that on P5: the affiliation on this publication P6 is I5 in country C4.

Say that this entity has received 3 citations from P1, P2 and P3.

Publication	Author	Institution	Country
P4	A2	I2	C2
	A5	I2	C2
P5	A1	I3	C2
		I1	C1
	A3	I2	C2
P6	A3	I5	C4
	A1	I5	C4
		I3	C2



P4, P5 and P6 are cited by

Publication	Author	Institution	Country
P1	A1	I6	C4
	A3	I6	C4
P2	A3	I2	C2
		I3	C2
	A4	I3	C2
P3	A4	I4	C3
	A5	I2	C2
		I6	C4
	A6	I6	C4
	A2	I4	C3

Question: What happens when the user chooses not to include self-citations when calculating a metric that offers this as an option?

Answer: If the entity you are viewing is, or includes, Countries C1, C2 or C4, then:

- A citation from P1 to P6 is a self-citation because they share the affiliation country C4
- A citation from P2 to all of P4, P5 and P6 is a self-citation: P2 shares affiliation country C2 with each of P4, P5 and P6
- A citation from P3 to all of P4, P5 and P6 is a self-citation: P3 shares affiliation country C2 with each of P4, P5 and P6, and in addition shares affiliation country C4 with P6
- A citation from P1 to P4 and P5 is not a self-citation because P1 does not share a country with either P4 or P5

Answer: If the entity you are viewing is, or includes, Institutions I1, I2, I3 or I5, then:

- A citation from P2 to all of P4, P5 and P6 is a self-citation: P2 shares institution I2 with both P4 and P5, and institution I3 with P5 and P6
- A citation from P3 to P4 and P5 is a self-citation: P3 shares institution I2 with both P4 and P5
- A citation from P1 to any of the entity's publications is not a self-citation because it does not share an institution with any of P4, P5 or P6
- A citation from P3 to P6 is not a self-citation because these publications do not have any institution in common

Answer: If the entity you are viewing is, or includes, authors A1, A2, A3 or A5, then:

- A citation from P1 to both P5 and P6 is a self-citation: P1 shares authors A1 and A3 with both of P5 and P6
- A citation from P2 to both P5 and P6 is a self-citation: P2 shares author A3 with both P5 and P6
- A citation from P3 to P4 is a self-citation: P3 shares authors A2 and A5 with P4
- A citation from P1 or P2 to P4 is not a self-citation because these publications do not have any author in common
- A citation from P3 to P5 or P6 is not a self-citation because these publications do not have any author in common

3. Selection of appropriate metrics

The ideal situation, when making research management decisions, is to have 3 types of input: peer review, expert opinion, and information from an evidence-base. When these complementary approaches “triangulate” to give similar messages, the user can have a higher degree of confidence that their decision is robust. Conflicting messages are a useful alert that further investigation is probably a good use of time.

It is also preferable that an evidence-base is used to illuminate a question from various angles. Multiple people are often asked to give their professional judgment about a question, and more than 1 peer review is typically sought; in just the same way, triangulating information about the same question from an evidence-base by using 2, 3 or even more different metrics will also ensure that the insights gained in this “corner of the triangle” are the most reliable they can be.

There are not really any strict rules about the selection of which metrics to use, besides approaching a question from more than 1 direction. The most appropriate metrics will always depend on the particular question that the user is asking. The best approach is to highlight some key points that are important to keep in mind, and for the user to apply their common sense.

SciVal offers a broad range of metrics to enable triangulation from the evidence-base, and to cater for the enormous variety of questions that users will ask. It is a rich and powerful resource of information, and can be used responsibly and appropriately by keeping a few facts in mind, as a complement to other sources of information. These facts are the focus of this section.

3.1 Clarity on the question being asked

The aim of using data and metrics as input into decision making is that any differences observed should reflect differences in performance. This will be the case if the user selects metrics that are suitable to answer their question, which in turn relies on 2 important factors:

- The question that is being asked is clearly articulated
- The user is aware of other factors, beyond performance, that can influence the value of a metric. These may or may not be important in the context of the questions being asked, but this judgment can only be made once that question has been clearly articulated

The types of questions asked typically fall into 3 groups:

- **Evaluation of performance**, such as is conducted by a national body on its research institutions for the purposes of allocating national funding, or by a line manager to provide input into career development discussions. It is typically very important in these situations that variables besides differences in performance have been accounted for to ensure that the assessment is fair; it would not be advisable to compare chemistry and immunology using metrics that do not take into account the tendency for higher output and citation rates in immunology, for instance
- **Demonstration of excellence**, such as that which may support an application for competitive funding, or that which may be used for promotional purposes to attract post-graduate students to a research institution. The aim in these situations is typically to find a way to showcase a particular entity, and a user may be able to benefit from the factors that affect a metric besides performance; for instance, a big institution may choose to use one of the “power metrics” that tend to increase as the entity gets bigger, whereas a small institution may choose to use a size-normalized metric
- **Scenario modeling**, such as that which supports the decision of which academic to recruit to an existing research team, or the thinking behind reorganizing a school. The importance of factors besides performance that affect the values of metrics may or may not be important, depending on the particular scenario that is being modeled

3.2 Factors besides performance that affect the value of a metric

Six types of factors, besides performance, that may affect the value of a metric are discussed in this section:

- Size
- Discipline
- Publication-type
- Database coverage
- Manipulation
- Time

The metrics available in SciVal inherently address some or none of these factors, as summarized in **Table 3**. In situations where a metric does not inherently account for a variable that may be important to address particular questions, SciVal supports the user by providing options in terms of functionality and other metrics; these options are outlined in the sections of this Guidebook dedicated to each individual metric.

	Size-normalized?	Field-normalized?	Publication-type-normalized?	Resistant to database coverage?	Difficult to manipulate?	Time-independent?
Scholarly Output						
Journal Count						
Journal Category Count						
Citation Count						
Cited Publications						
Citations per Publication						
Number of Citing Countries						
Field-Weighted Citation Impact						
Collaboration						
Collaboration Impact						
Academic-Corporate Collaboration						
Academic-Corporate Collaboration Impact						
Outputs in Top Percentiles						
Publications in Top Journal Percentiles						
<i>h</i> -indices						

Table 3: Characteristics of the metrics in SciVal. Those metrics that have half a shaded cell in the “Size-normalized” column are size-normalized when the “Percentage” option is selected, but not when the “Absolute Number” option is selected.

3.2.1 Size

There are some metrics whose value tends to increase with the size of an entity, such as Scholarly Output that indicates the productivity of an entity, and Citation Count that sums, over all publications, the citations received by an entity. These metrics are referred to within this Guidebook as “Power metrics”, and are summarized in [Table 1](#).

It is often important to account for differences in the size of an entity when evaluating performance; Citations per Publication, for instance, accounts for differences in the size of an entity’s Scholarly Output, and is useful to reveal efficiency of citations received per item. When demonstrating excellence, however, “Power metrics” may be used to the benefit of large entities whose performance will tend to look better; Citation Count will generally be higher for a large collaboration network than for a small research team.

3.2.2 Discipline

Academics working in different disciplines display distinct characteristics in their approach to research and in their communication about research. These behavioral differences are not better or worse than each other, but are merely a fact associated with particular fields of research.

Any Citation Count-type or Citations per Publication-type metric effectively illustrates these differences. These metrics tend to be significantly higher in neuroscience than in engineering, for example, but it is obviously not the case that neuroscience is generally “better” than engineering. These types of metric do not take into account different behaviors between fields, and it is not advised to use them to make comparisons between fields: for this purpose, field-normalized metrics, such as Field-Weighted Citation Impact, Publications in Top Journal Percentiles, and the free journal metrics SNIP and SJR¹², are suitable.

It is not only Citation Count-type and Citations per Publication-type metrics that are affected by these behavioral differences; they affect all metrics used throughout research management, including those implemented in SciVal. What is it that causes these differences?

- **Frequency of publication.** Academics in fields such as chemical engineering tend to publish more often than researchers in mathematics
- **Length of reference lists.** Publications in disciplines such as toxicology tend to have much longer reference lists than those in social sciences
- **Number of co-authors.** Research in physics tends to be much more collaborative than research in arts and humanities, resulting in a higher number of co-authors per publication

The typical distribution of these behaviors amongst all disciplines is illustrated in **Figure 1**.

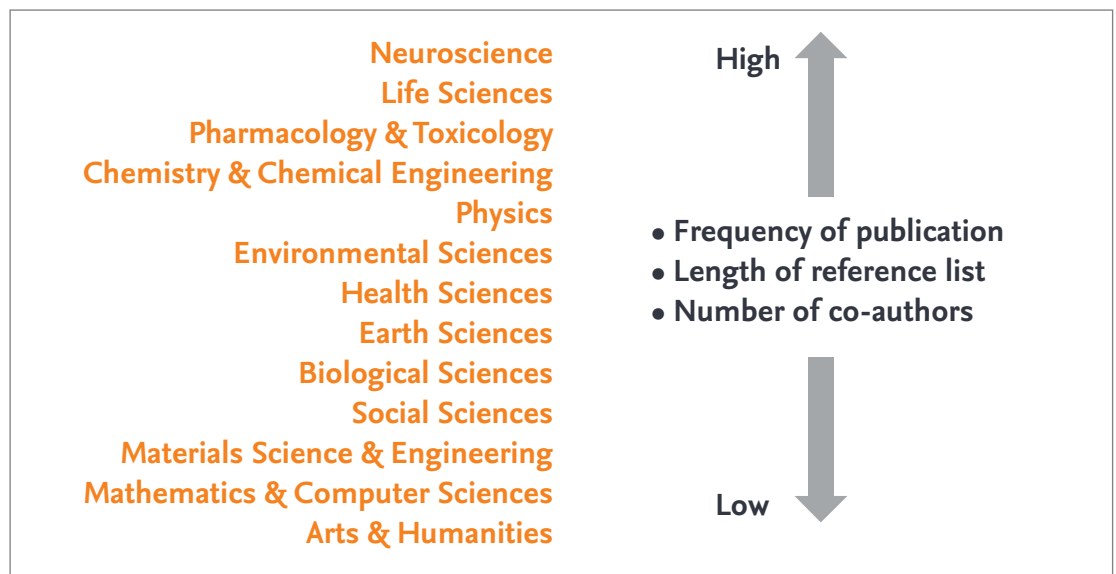


Figure 1: The characteristic behavior of academics differs between disciplines

12: www.journalmetrics.com

What does this mean for metrics? The majority of metrics values tend to be higher in neuroscience and life sciences than they are in materials science and computer sciences; this does not necessarily reflect performance, but merely the way that academics in these disciplines conduct research. This may be used to advantage when the aim is to demonstrate excellent performance of an academic working in neuroscience, for instance, but it is very important to be aware of this effect and to take it into account when performing an evaluation.

If users need to make comparisons between disciplines, they are advised to:

- **Apply the Research Area filter** when comparing entities made up of a mixture of disciplines, such as an Institution or a Country, to focus on one field that is common between all the entities
- **Select field-normalized metrics**, such as Field-Weighted Citation Impact

It is worth noting, when considering the effect of disciplines on metric values, that there are different ways of defining disciplines:

- **Journal classifications** are probably the most common approach to defining disciplines
 - Scopus, for example, assigns the journals that it indexes into Main Categories and Sub-Categories, and the Brazilian CAPES Foundation¹³, a government agency that awards scholarship grants to graduate students at universities and research centers in Brazil and abroad, also has a journal classification, as do many other organizations
 - These classifications offer the significant advantage to users of a tool like SciVal of having the same meaning, regardless of the entity being investigated. The meaning of a chemistry department may differ between research institutions, and that is a concern when benchmarking chemistry. If a user filters institutions by a particular journal category they can be sure that they are comparing a consistent definition of a discipline. It is for this reason that the Research Area filter in SciVal offers journal classifications
 - The drawback of these classifications is that they tend to be very large so that it is difficult to keep them up to date with current developments in research, and that it is assumed that a journal's discipline(s) applies equally to each individual publication that it contains

13: <http://www.iie.org/en/Programs/CAPES>

-
- **Publication-level classifications** are gaining increased attention, as the pace of change in academia increases and with technological advances that can handle massive data sets:
 - These classifications do not assume that a publication fits into the same discipline(s) as the journal that it appears in. They are often formed by allowing publications to group themselves through the references they contain, and/or by the publications that refer to them
 - Publication-level classifications are very useful to keep pace with the changing ways in which academics view the world of research, and can be useful to reveal nascent, emerging fields. The global research strengths displayed in the competency maps of SciVal are examples of research activity that represent the way that recent publications cite previous work. It is also for this reason that publication-level classifications are used to determine disciplines in the calculation of Field-Weighted Citation Impact in SciVal
 - The drawbacks of these classifications are:
 - There tend to be publications which cannot be easily assigned, for instance if they have a short or absent reference list and/or have not yet received citations
 - It is computationally very demanding to implement an ever changing classification
 - Frequent changes in classification due to changes in the pattern of citations received reduce the transparency of the data underlying a metric calculation for users
 - **User-defined disciplines**, for instance by searching for key phrases within the titles and abstracts of a publications database:
 - The advantage of user-defined disciplines is that they ensure that every user can define the world of research as it makes sense to them and their questions. Academics do not conduct their research in a way that conveniently fits the journal classifications and previous citing behavior of their colleagues, and even publication-level classifications are not able to detect the very first stages of Research Areas, when academics are just starting to focus on a new field
 - The drawback of user-defined disciplines is that they are likely unique, and pose challenges to others who would like to use or validate these customized fields

3.2.3 Publication-type

Different types of publications tend to be cited with different rates. The most well-known example is that reviews tend to attract more citations than original research articles, but there are variations between other publication-types as well. These are illustrated in **Figure 2**.

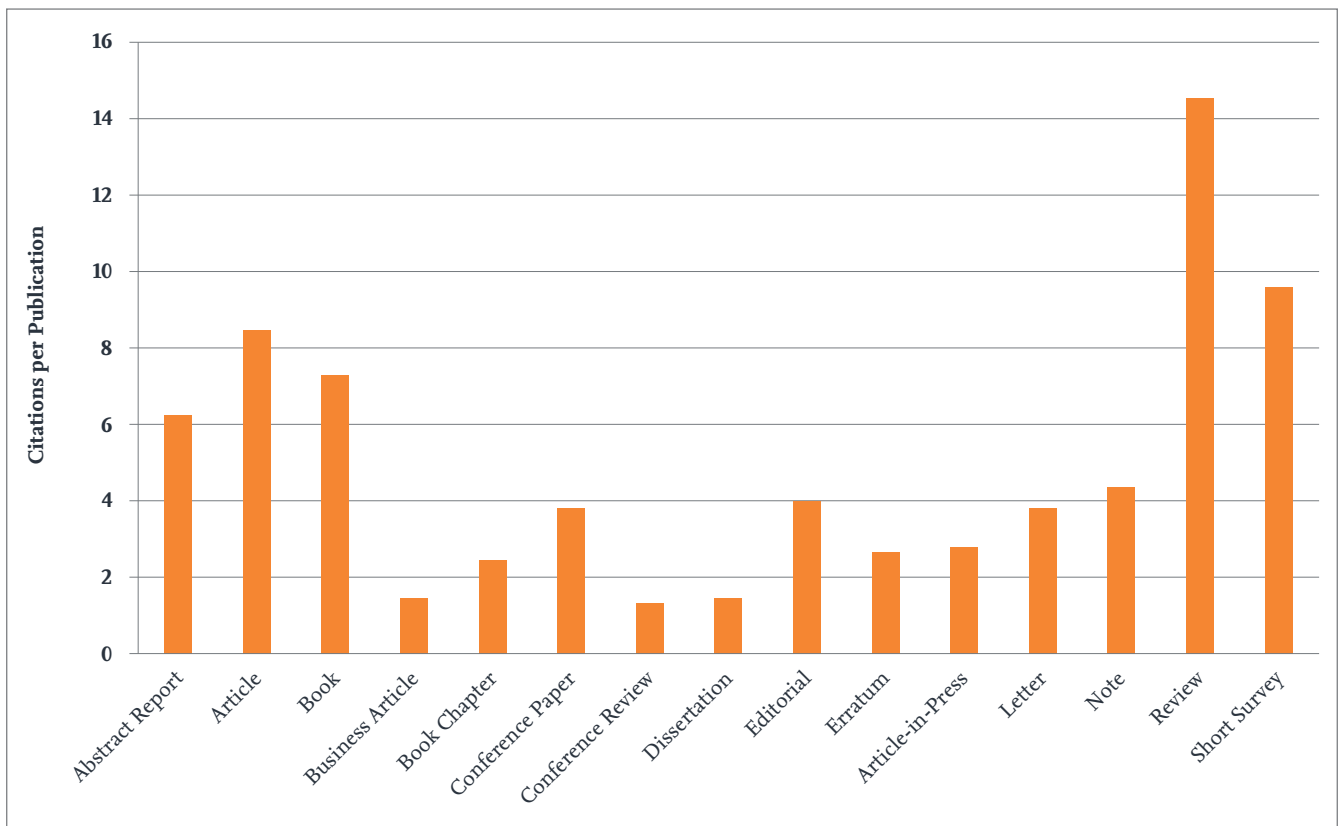


Figure 2: Citation rates for different publication-types as classified in Scopus. This chart displays citations received up to August 2013 per item published during the period 2008-2012.

Publication-type differences may not be important where the performance of total output is of interest, or when a journal editor is showcasing their productivity and wishes to retain the positive effect of editorials in the count. There are some situations, however, when there may be a strong preference to focus on a particular type, such as on conference proceedings in engineering and computer science, or original research articles when comparing a journal editor who has written many editorials with an academic who is not a journal editor.

If this effect is important for the question being addressed by the user, they are advised to:

- **Apply the Publication-type filter**
- **Use Field-Weighted Citation Impact** which is normalized for Publication-type

3.2.4 Database coverage

Databases have particular guidelines in determining which content to include. Scopus has a comprehensive policy¹⁴ to select the content which meets its aims, and the policy is to be selective and not to include every single publication globally. This means that there may be some items that have been published by a particular entity that are not indexed in Scopus, and so cannot be part of the metrics calculations in SciVal.

There are 2 aspects to considerations of database coverage:

- **Geographical coverage.** Scopus indexes content from more than 5,000 publishers from all over the world. The geographical distribution of titles indexed in Scopus is representative of the global concentrations of publishers, with the focus of activity in the United States and the United Kingdom, as shown in **Figure 3**. This geographical coverage should support a thorough analysis of topics of global interest; however, for research areas of primarily local interest, such as national literature, history or culture, Scopus may not provide sufficiently complete data
- **Disciplinary coverage.** The ongoing expansion of the titles indexed by Scopus means that this coverage will continue to change. The disciplinary coverage of Scopus can be estimated by looking at the items that have been cited by recently published work; the extent to which these citations can be linked to items indexed within the Scopus database represents the coverage, and those citations which refer to items not indexed by Scopus are assumed to represent lack of coverage. This methodology represents a minimum coverage level, and the actual coverage is probably a few percentages higher than shown in **Figure 4**

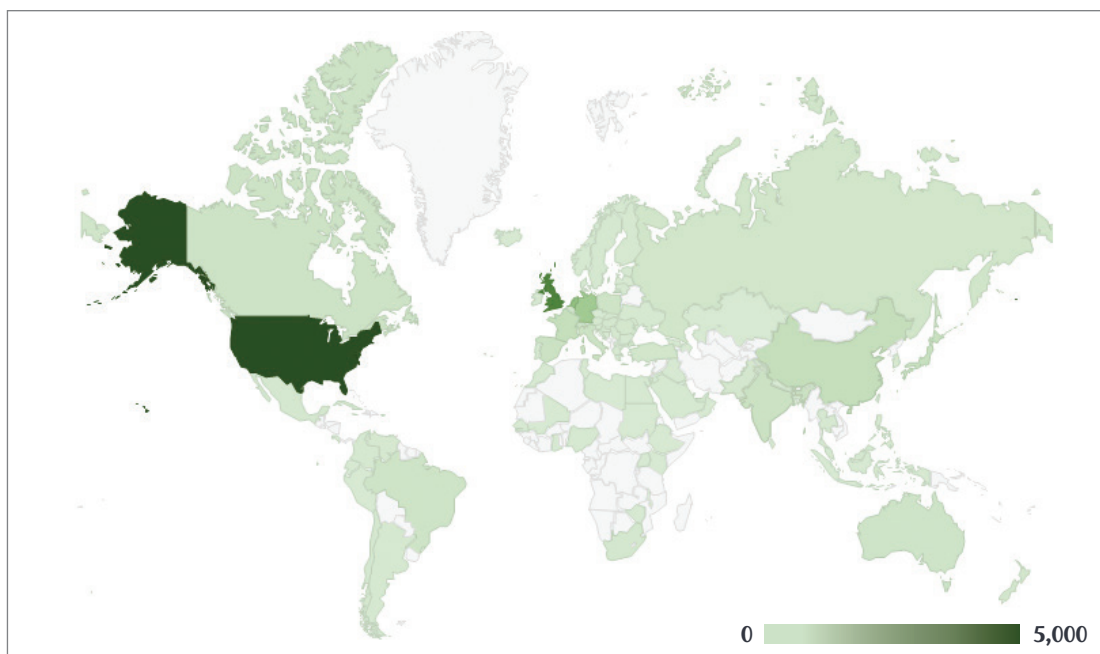


Figure 3: Geographical distribution of titles indexed in Scopus, to January 2013. The country assigned to a title is the country of the publisher imprint. The sliding scale of 0-5,000 indicates the density of indexed titles.

14: <http://www.elsevier.com/online-tools/scopus/content-overview>

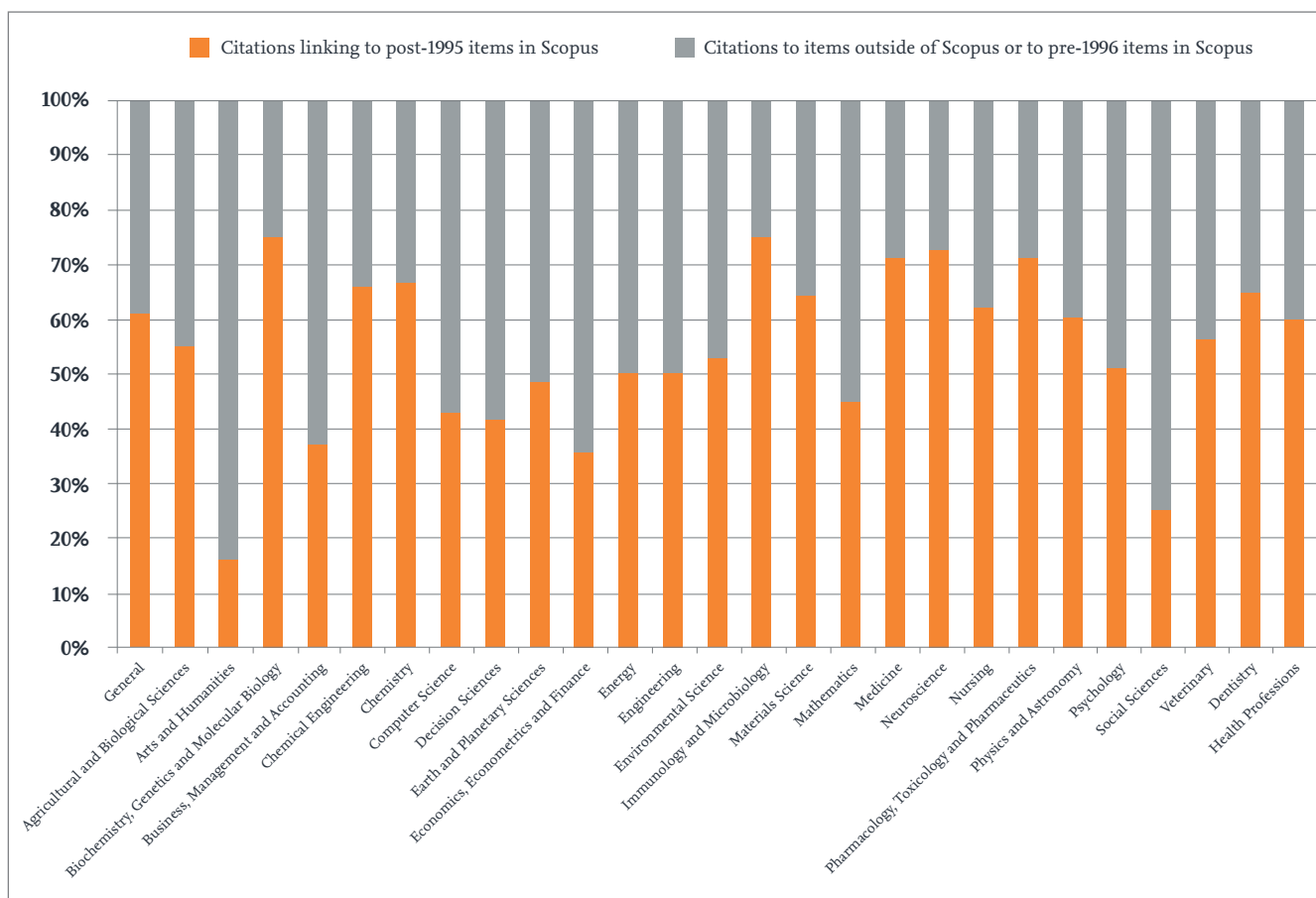


Figure 4: Estimation of Scopus' disciplinary coverage. This estimation is made based on the extent to which citations from publications during the period 2008-2012 can be linked to items indexed from 1996 onwards within the Scopus database; this is an under-estimation of the true coverage because citations to indexed items published before 1996 are not captured here. This analysis is based on data up to August 2013.

It is most probably acceptable to have a few missing publications from an entity's output when benchmarking large data sets against their peers, because gaps in database coverage likely affect all entities similarly and will not invalidate the comparison. Care is advised, however, when comparing small entities, from which a single missing publication may have a significant negative impact; for example, an academic's performance may suffer due to gaps in database coverage of their portfolio, as well as gaps in publications citing those items of the portfolio that are indexed. The transition point between a "small" and a "large" entity is a matter of judgment, and will differ depending on the discipline. The only way to account for this is to be vigilant and to apply common sense when using SciVal to support decisions.

The question of the effect of database coverage is most often raised in relation to **Arts and Humanities, and Social Sciences**: are the metrics in SciVal useful in these fields? In some situations, they are:

- The type of decisions supported by SciVal should ideally be based on a combination of inputs from peer review and expert opinion, as well as from an evidence base. Publication and citation data form one part of this evidence-base, with funding, innovation, and societal impact, for instance, also being very important although they are not yet addressed in SciVal. This is the case for all disciplines, and Arts and Humanities and Social Sciences are no exception
- There is a further concern with these fields, in that their coverage in Scopus and other commercial databases tends to be low. This is a natural consequence of the publishing behavior of academics in these fields, who favor the publication of stand-alone books which are difficult to capture relative to serials, although Scopus is now focusing on increasing its coverage of these items. Nevertheless, valuable information about performance in these fields can be gained from SciVal if the guidelines about the size of the entity, above, are borne in mind

3.2.5 Manipulation

Some situations lend themselves relatively easily to manipulation for the artificial inflation of metrics values. One example is the combination of research units to artificially increase size for reporting purposes, which tends to improve apparent performance when using “Power metrics”.

Another example is that of self-citations. There is nothing inherently wrong about self-citing: it is normal academic behavior to build on work previously published, and it is an author's responsibility to refer a reader to older material that will support their understanding. This practice is, however, open to abuse by unscrupulous academics who could choose to cite irrelevant past papers to boost their citation counts, and to journal editors who occasionally coerce authors of submitted papers to add too many additional citations to publications within their journal to their reference lists.

Citation Impact metrics are the most susceptible to manipulation of the metrics in SciVal. Self-citation abuse is very rare, but users have the option to exclude self-citations from several of these metrics if they wish, as illustrated in **Example 1**. Other metrics are much more difficult to manipulate, such as those that provide information on aspects of productivity and collaboration.

3.2.6 Time

The passage of time is critical to enable useful information to be derived from some metrics. The most obvious of these are Citation Impact metrics, since time is needed for published work to receive citations. The *h*-indices are another example of metrics that are not very useful when applied to the output of an early-career researcher, because the need for the passage of time to accumulate citations is coupled with that for a reasonably-sized body of work to provide reliable information.

Some “time-independent” metrics can be used immediately upon entry of a publication into the database. Examples are metrics that look at collaboration by using affiliation information. There is even a Citation Impact metric that is independent of time; Publications in Top Journal Percentiles uses citations received by the journals in which publications appear.

4. SciVal metrics: methods and use

This section covers the individual metrics available in SciVal. It shares their method of calculation, situations in which they are useful, and highlights situations in which care should be taken when using them.

It also suggests metrics that might be considered useful partners, either to address shortcomings of a particular metric, or to highlight information that is naturally complementary. These suggestions are summarized in **Table 4**; they should not be taken as rules, but as guidelines that may sometimes be useful to the user in their selection of appropriate metrics.

...useful partner metrics are:

	Scholarly Output	Journal Count	Journal Category Count	Citation Count	Cited Publications	Citations per Publication
Scholarly Output						
Journal Count						
Journal Category Count						
Citation Count						
Cited Publications						
Citations per Publication						
Number of Citing Countries						
Field-Weighted Citation Impact						
Collaboration						
Collaboration Impact						
Academic-Corporate Collaboration						
Academic-Corporate Collaboration Impact						
Outputs in Top Percentiles						
Publications in Top Journal Percentiles						
<i>h</i> -indices						

Table 4: Suggested partner metrics

...useful partner metrics are:

[illegible]

- | |
|---|
| “Power metrics” |
| Time-independent metrics |
| Communicate information about the magnitude of metrics values |
| Natural complementary partner metrics |
| Rely on actual performance of publications, not journal average |
| Reliability with which a publication will receive at least 1 citation |
| Avoid display of the “citation dip” in recent years |

4.1 The display of metrics in SciVal

The metrics in SciVal are displayed exactly as calculated. There may be situations in which users would prefer values to be averaged over a multiple year period to “smoothen” fluctuations in trend lines, but such “smoothening” is not performed in SciVal so that the data underlying a metric are very transparent.

This means that some trend lines in SciVal will look “jumpy”, especially when they are calculated on a small data set, such as a Researcher. Trend lines tend to become smoother when metrics for larger data sets are being displayed.

4.2 Metric: Scholarly Output ❄️

Scholarly Output in SciVal indicates the productivity of an entity: how many publications does this entity have indexed in Scopus?

Scholarly Output is a:

- Productivity metric
- Snowball Metric
- “Power metric”

This metric is useful to:

- **Benchmark the productivity of entities that are similar** in size and that fall within similar fields of research, such as Institutions with a similar number of research staff, or Researchers in the same field of research and with similar career lengths
- Provide impressive figures to **showcase the performance of entities that are large** in comparison to a group of peers, when this metric is likely to give high numbers
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal
- Look at publishing activity in a way that is **difficult to manipulate**
- **Investigate activity during the early stages** of a new strategy when publications are just starting to appear, or of early-career researchers

This metric should be used with care when:

- **Benchmarking the productivity of entities of obviously different sizes**, such as Countries and Institutions, large collaboration networks and individual Researchers, or stable, established Research Areas and small, emerging Research Areas:
 - Differences in the value of this “Power metric” for comparing such entities will probably reflect distinct entity sizes rather than differences in productivity
 - Users are advised to limit the use of this metric to similar entities, or to use size-normalized metrics such as Citations per Publication
- **Benchmarking the collaboration of entities with distinct disciplinary profiles:**
 - Academics in distinct fields tend to publish with very different frequencies. For example, Institutions with a high proportion of academics in life sciences, who tend to publish with high frequency, are likely to have a higher Scholarly Output than similarly sized Institutions with a high proportion of academics in humanities, who tend to publish with low frequency
 - Differences in Scholarly Output, in these cases, are most likely to reflect distinct disciplinary characteristics rather than to provide reliable information about differences in productivity

-
- It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, users are advised to apply the Research Area filter to focus on one field that is common between all the entities, or to use field-normalized metrics such as Publications in Top Journal Percentiles which will take this into account
- **Understanding the productivity of small entities for which there may be gaps within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on apparent productivity
 - The only way to account for this is to be vigilant. Consider also limiting the use of Scholarly Output to comparing larger data sets where gaps in the database coverage likely have a similar effect on all entities and do not invalidate the comparison

Useful partner metrics are:

- **Citation Count**, which sums, over all publications, the citations received by an entity, and is a natural complement to Scholarly Output that counts the publications of an entity
- ***h*-indices**, whose values depend on a combination of Scholarly Output together with Citation Count, and are natural partner metrics
- **The set of all other “Power metrics”** whose value tends to increase as the entity becomes bigger: Journal Count, Journal Category Count, Citation Count, Cited Publications (“Absolute number”), Number of Citing Countries, Collaboration (“Absolute number”), Academic-Corporate Collaboration (“Absolute number”), Outputs in Top Percentiles (“Absolute number”), Publications in Top Journal Percentiles (“Absolute number”), and *h*-indices
- **The set of all other “time-independent metrics”** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 2**.

4.3 Metric: Journal Count

Journal Count in SciVal indicates the diversity of an entity's publication portfolio: in how many of the distinct journals indexed in Scopus have this entity's publications appeared?

Journal Count is a:

- Disciplinarity metric
- “Power metric”

This metric is useful to:

- **Compare smaller entities**, such as Researchers and Publication Sets, where differences in the journal portfolio are most likely to be evident
- Provide impressive figures to **showcase the performance of a relatively large entity**, when this metric is likely to give high numbers
- **Benchmark the diversity of the publication portfolios of related entities**, such as scenario models of a research institute created to investigate the effect of recruiting various researchers, or collaboration networks in a given field of research
- **Demonstrate the excellence of entities that work across traditional disciplines** and have a broad array of journals available to which to submit
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal
- Look at publishing activity in a way that is **difficult to manipulate**
- **Investigate differences in activity at the early stages** of a new strategy when publications are just starting to appear, or between early-career researchers

This metric should be used with care when:

- **Comparing large entities**, such as Institutions and Groups of Institutions, which will likely publish in such a large and stable journal portfolio that it will approach the maximum extent of the Scopus database so that differences may not be visible. Users are advised to narrow their focus to a slice of a large entity when using this metric, such as by applying the Research Area filter
- **Benchmarking the productivity of entities of obviously different sizes**, such as departments of 150 academics with departments of 30 academics:
 - Differences in the value of this “Power metric” when comparing such entities will probably reflect distinct entity sizes rather than differences in the diversity of their publication portfolios
 - Users are advised to limit the use of this metric to similar entities, or to use size-normalized metrics such as Collaboration

-
- **Benchmarking the productivity of entities with distinct disciplinary profiles**, when academics in distinct fields may have a different sized range of journals available for submission:
 - Differences in Journal Count, in these cases, may well reflect distinct disciplinary characteristics rather than provide reliable information about differences in publication portfolios
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, users are advised to apply the Research Area filter to focus on one field that is common between all the entities, or to use field-normalized metrics such as Field-Weighted Citation Impact which will take this into account
 - **Understanding the productivity of small entities for which there may be gaps within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on the apparent breadth of a publication portfolio
 - The only way to account for this is to be vigilant. Consider also limiting the use of Journal Count to comparing slightly larger data sets where gaps in the database coverage likely have a similar effect on all entities and do not invalidate the comparison

Useful partner metrics are:

- **Journal Category Count**, which highlights the disciplinary portfolio of an entity, and is a natural complement to Journal Count
- **The set of all other “Power metrics”** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Category Count, Citation Count, Cited Publications (“Absolute number”), Number of Citing Countries, Collaboration (“Absolute number”), Academic-Corporate Collaboration (“Absolute number”), Outputs in Top Percentiles (“Absolute number”), Publications in Top Journal Percentiles (“Absolute number”), and *h*-indices
- **The set of all other “time-independent metrics”** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 2**.

4.4 Metric: Journal Category Count

Journal Category Count in SciVal indicates the diversity of an entity's disciplinary portfolio: in how many distinct journals categories have this entity's publications appeared?

Journal Category Count can be generated using either the Main Categories or Sub-categories. The maximum value of this metric is 27 when using Scopus Main Categories, and 334 when using Scopus Sub-categories. Publications can be assigned to a classification system in 2 ways:

- **“Journal-driven” assignment** assumes that every publication within a journal fits within the same discipline(s) as the journal's scope. Each publication automatically adopts the subject classifications that are assigned to the journal. This method of assignment is suitable for journals that are focused in a core field, and do not tend to include publications that are also a relevant to other fields
- **“Publication-driven” assignment** assumes that publications within a journal may have additional or different relevance to fields outside the core focus of the journal's scope. Publication-driven assignment offers the benefit of being able to assign individual publications from a journal separately to their relevant classifications. This is important for publications in multi-disciplinary journals

Journal Category Count uses “publication-driven” assignment, and a publication can be allocated to more than 1 category.

Journal Category Count is a:

- Disciplinarity metric
- “Power metric”

This metric is useful to:

- **Compare smaller entities**, such as Groups of Researchers and Groups of Publication Sets, where differences in the disciplinary portfolio are most likely to be evident
- Provide impressive figures to **showcase the performance of a relatively large entity**, when this metric is likely to give high numbers
- **Benchmark the diversity of the disciplinary portfolios of related entities**, such as collaboration networks funded by the same type of grant from a given funder
- **Provide evidence of cross-disciplinarity** by indicating the appeal of an entity's output to readers in diverse disciplines
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal
- Look at publishing activity in a way that is **difficult to manipulate**
- **Investigate differences in activity at the early stages** of a new strategy when publications are just starting to appear, or between early-career researchers

This metric should be used with care when:

- **Comparing large entities**, such as Institutions and Groups of Countries, which will likely publish in such a large and stable disciplinary portfolio that it will approach the maximum extent of the Scopus database and differences may not be visible. Users are advised to narrow their focus to a slice of a large entity when using this metric, such as by applying the Research Area filter
- **Benchmarking the productivity of entities of obviously different sizes**, such as large collaboration networks with single Publication Sets:
 - Differences in the value of this “Power metric” when comparing such entities will probably reflect distinct entity sizes rather than differences in the diversity of their disciplinary portfolios
 - Users are advised to limit the use of this metric to similar entities, or to use size-normalized metrics such as Academic-Corporate Collaboration

-
- **Benchmarking the productivity of entities with distinct disciplinary profiles**, when academics in distinct fields may have different opportunities to work in a cross-disciplinary manner:
 - Differences in Journal Category Count, in these cases, may well reflect distinct disciplinary characteristics rather than provide reliable information about differences in disciplinary portfolios
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, users are advised to apply the Research Area filter to focus on one field that is common between all the entities, or to use field-normalized metrics such as Publications in Top Journal Percentiles which will take this into account
 - **Understanding the productivity of small entities for which there may be gaps within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on the apparent breadth of a disciplinary portfolio
 - The only way to account for this is to be vigilant. Consider also limiting the use of Journal Category Count to comparing slightly larger data sets where gaps in the database coverage likely have a similar effect on all entities and do not invalidate the comparison

Useful partner metrics are:

- **Journal Count**, which highlights the publication portfolio of an entity, and is a natural complement to Journal Category Count
- **The set of all other “Power metrics”** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Citation Count, Cited Publications (“Absolute number”), Number of Citing Countries, Collaboration (“Absolute number”), Academic-Corporate Collaboration (“Absolute number”), Outputs in Top Percentiles (“Absolute number”), Publications in Top Journal Percentiles (“Absolute number”), and *h*-indices
- **The set of all other “time-independent metrics”** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Collaboration, Academic-Corporate Collaboration, and Publication in Top Journal Percentiles

See **Example 2**.

Example 2: Scholarly Output, Journal Count and Journal Category Count

Scenario: The user would like to calculate the Scholarly Output, Journal Count, or Journal Category Count of an entity that consists of 6 publications, and has selected the following viewing and calculation options.

Selected Publication Year range	2008 to 2012
Selected Publication Types	Articles, reviews and conference papers
Selected Research Area	Agricultural and Biological Sciences

Entity with 6 Publications				
Publication Identity	Publication 1	Publication 2		
Publication Year	2007	2009		
Journal in which publication is published	<i>Bioscience Journal</i>	<i>Biology and Environment</i>		
Publication Type	Article	Article		
Journal Sub-Category (or Sub-Categories)	General Agricultural and Biological Sciences	Animal Science and Zoology	Aquatic Science	
Journal Main Category (or Categories)	Agricultural and Biological Sciences	Agricultural and Biological Sciences	Agricultural and Biological Sciences	
Do publications match user-selected options?				
Step 1	Does the Journal Category of the publication match the selected Research Area?	Yes	Yes	Yes
Step 2	Does this publication fall in the selected Publication Year range?	No	Yes	
Step 3	Does the Publication Type match the selected Publication Types?	Yes	Yes	
Step 4	Does the publication pass each of steps 1, 2 and 3?	No	Yes	
Question: How do I calculate Scholarly Output?	Answer: Count the number of publications that received a “yes” in step 4.	Scholarly Output = 3		
Question: How do I calculate Journal Count?	Answer: Note unique Journal Titles for those publications that received a “yes” in step 4. – Count the number of unique Journal Titles.	Remove duplicates Journal Titles of those publications that passed the filters. <i>Biology and Environment, Archiv für Lebensmittelhygiene</i> Journal Count = 2		
Question: How do I calculate Journal Category Count (Main Categories)?	Answer: Look up main categories. – Remove duplicates. – Count the number of unique Main Categories.	Remove duplicate Main Category names from those publications that passed the filters. <i>Agricultural and Biological Sciences</i> Journal Main Category Count = 1		
Question: How do I calculate Journal Category Count (Sub-categories)?	Answer: Look up sub-categories. – Remove duplicates. – Count the number of unique Sub-Categories.	Remove duplicate Sub-Category names from those publications that passed the filters. <i>Agronomy and Crop Science, Animal Science and Zoology, Aquatic Science, Horticulture</i> Journal Sub-Category Count = 4		

Entity with 6 Publications				
Publication 3		Publication 4	Publication 5	Publication 6
2009		2008	2010	2010
<i>Biology and Environment</i>		<i>Biology and Environment</i>	<i>Archiv für Lebensmittelhygiene</i>	<i>Archiv für Lebensmittelhygiene</i>
Article In Press		Review	Article	Report
Insect Science	Environmental Chemistry	Agronomy and Crop Science	Horticulture	Organic Chemistry
Agricultural and Biological Sciences	Environmental Science	Agricultural and Biological Sciences	Agricultural and Biological Sciences	Chemistry
Yes No		Yes	Yes	No
Yes		Yes	Yes	Yes
No		Yes	Yes	No
No		Yes	Yes	No

4.5 Metric: Citation Count

Citation Count in SciVal indicates the total citation impact of an entity: how many citations have this entity's publications received?

Citation Count is a:

- Citation Impact metric
- Snowball Metric
- “Power metric”

SciVal often displays Citation Count in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Benchmark the visibility of entities of similar size**, and that fall into similar disciplines, such as multidisciplinary institutions with a similar number of research staff, or international collaboration networks in similar disciplines
- Provide impressive figures to **showcase the performance of entities that are large** in comparison to a group of peers, when this metric is likely to give high numbers
- **Showcase the performance of entities that have published a few noticeably highly cited papers**
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

This metric should be used with care when:

- **Benchmarking the visibility of entities of obviously different sizes**, when this “Power metric” will most closely reflect entity size rather than differences in visibility; a Group of Institutions such as a US state will generally have a higher Citation Count than an individual institution, for example. Users are advised to use the size-normalized Citation Impact metrics Citations per Publication or Field-Weighted Citation Impact to compare the visibility of entities of different sizes

-
- **Benchmarking the collaboration of entities with distinct disciplinary profiles:**
 - Academics working in medical sciences and in virology, for instance, tend to publish and cite with high frequencies, whereas those working in business schools or in linguistics tend to publish and cite with lower frequencies
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities, or to select Field-Weighted Citation Impact which will take this into account
 - **Investigating the reliability with which an entity's publications are cited**, since one or a few publications with a very high number of citations can conceal a sizeable body of uncited material. Users are advised to select Cited Publications to give an idea of reliability
 - **Entities are small and there may be gaps in their output within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on apparent visibility. For example, an academic's performance may suffer due to gaps in items they have published, as well as gaps in publications citing the publications that are indexed
 - The only way to account for this is to be vigilant. Consider also limiting the use of Citation Count to comparing larger data sets in the same discipline where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison
 - There is a **concern that excessive self-citations may be artificially inflating the number of citations**. Users can judge whether the level of self-citations is higher than normal by deselecting the "Include self-citations" option
 - **Uncovering the performance of publications in the very early stages of a new strategy, or of early-career researchers**, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. Users are advised to use metrics such as Scholarly Output or Collaboration in these situations
 - **The person who will use the data does not like to see a line that "dips" in recent years**. This typically happens with Citation Count because recent publications have had little time to receive citations. Users are advised to use Field-Weighted Citation Impact or Publications in Top Journal Percentiles to avoid this drop, if it is of concern

Useful partner metrics are:

- **Citations per Publication and Field-Weighted Citation Impact**, which bring complementary perspectives into view on total visibility, and also account for differences in the size of entities being compared. Field-Weighted Citation Impact also accounts for differences in publication and citation behavior between disciplines
- **Field-Weighted Citation Impact, Outputs in Top Percentiles, and Publications in Top Journal Percentiles** avoid the “dip” in recent years due to little time having passed to receive citations since publication
- **Cited Publications** provides a measure of the reliability that an entity’s publications will subsequently be cited, and is not affected by a high number of citations received by 1 or a few publications
- ***h*-indices**, whose values depend on a combination of Citation Count together with Scholarly Output, and are natural partner metrics
- **The set of all other “Power metrics”** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Cited Publications (“Absolute number”), Number of Citing Countries, Collaboration (“Absolute number”), Academic-Corporate Collaboration (“Absolute number”), Outputs in Top Percentiles (“Absolute number”), Publications in Top Journal Percentiles (“Absolute number”), and *h*-indices
- **The set of “time-independent metrics”** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 3**.

4.6 Metric: Cited Publications

Cited Publications in SciVal indicates the citability of a set of publications: how many of this entity's publications have received at least 1 citation?

Cited Publications is a:

- Citation Impact metric
- “Power metric” when the “Absolute number” option is selected, but not when the “Percentage” option is selected

SciVal often displays Cited Publications in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Benchmark the extent to which publications are built on by subsequent work**
- **Compare the influence of publications of entities of different sizes**, but in related disciplines, such as large and small countries, or a research team with individual Researchers within that team
 - It is advised to select the “Percentage” option when comparing entities of different sizes, to normalize for this variable
- **Demonstrate the excellence of entities that produce consistent work that is reliably cited**, regardless of the number of citations received
- **Account for the positive impact of a few very highly cited papers** on apparent performance, which affect Citation Count and Citations per Publication
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

This metric should be used with care when:

- **Benchmarking the extent to which the publications of entities with distinct disciplinary profiles are built on by subsequent work:**
 - Teams working in parasitology, for instance, may experience a relatively short delay between publishing and the receipt of citations because of the high frequency of publishing and citing, relative to teams working in mathematical modeling where publication behavior may lead to a relatively long delay between publishing and the receipt of citations
 - It is not advisable to use this metric to compare entities in distinct disciplines without taking these differences into account
 - When comparing entities made up of a mixture of disciplines, such as an Institution or a Group of Institutions, it is advised to apply the Research Area filter to focus on one field that is common between all the entities

-
- **Understanding the magnitude of the number of citations received** by an entity's publications. Users are advised to use Citation Count or Citations per Publication to communicate this information
 - **Entities are small and there may be gaps in their output within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on apparent visibility. For example, an academic's performance may suffer due to gaps in items they have published, as well as gaps in publications citing the publications that are indexed
 - The only way to account for this is to be vigilant. Consider also limiting the use of Cited Publications to comparing larger data sets in the same discipline where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison
 - There is a **concern that excessive self-citations may be artificially inflating the proportion of cited publications**. Users can judge whether the level of self-citations is higher than normal by deselecting the "Include self-citations" option
 - **Uncovering the performance of publications in the very early stages of a new strategy, or of early-career researchers**, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information

Useful partner metrics are:

- **Citations Count and Citations per Publication**, which indicate the magnitude of citation impact, and which will be positively impacted by one or a few publications that have received a very high number of citations
- **The set of all other "Power metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Citation Count, Number of Citing Countries, Collaboration ("Absolute number"), Academic-Corporate Collaboration ("Absolute number"), Outputs in Top Percentiles ("Absolute number"), Publications in Top Journal Percentiles ("Absolute number"), and *h*-indices
- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 3**.

4.7 Metric: Citations per Publication ❄️

Citations per Publication in SciVal indicates the average citation impact of each of an entity's publications: how many citations have this entity's publications received on average?

Citations per Publication is a:

- Citation Impact metric
- Snowball Metric

SciVal often displays Citations per Publication in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Benchmark the average citation impact of publications within a body of work**
- **Compare the average influence of publications of entities of different sizes**, but in related disciplines, such as research teams working in a similar field of research, or a Researcher and Publication Sets belonging to that Researcher
- **Showcase the performance of entities that have published a few notably highly cited papers** that will have a positive effect on the average of the entire data set
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

This metric should be used with care when:

- **Benchmarking the average influence of the publications of entities with distinct disciplinary profiles,** such as Institutions with sizeable humanities schools with Institutions without humanities schools
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities, or to select Field-Weighted Citation Impact which will account for disciplinary differences
- **Investigating the reliability with which an entity's publications are cited,** since one or a few publications with a very high number of citations can conceal a sizeable body of uncited material. Users are advised to use Cited Publications to investigate the proportion of publications in a data set that have been cited
- **Entities are small and there may be gaps in their output within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on apparent visibility. For example, an academic's performance may suffer due to gaps in items they have published, as well as gaps in items citing the publications that are indexed
 - The only way to account for this is to be vigilant. Consider also limiting the use of Citations per Publication to comparing larger data sets in the same discipline where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison
- **Entities are small so that the metric may fluctuate significantly and appear unstable** over time, even when there is complete Scopus coverage. Citations per Publication calculates an average value, and these types of calculations are strongly influenced by outlying publications in a small data set
- There is a **concern that excessive self-citations may be artificially inflating Citations per Publication.** Users can judge whether the level of self-citations is higher than normal by deselecting the "Include self-citations" option
- **Uncovering the performance of publications in the very early stages of a new strategy, or of early-career researchers,** where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. It is advised to use metrics like Journal Category Count or Collaboration to account for this
- **The person who will use the data does not like to see a line that "dips" in recent years.** This typically happens with Citations per Publication because recent publications have had little time to receive citations. Users are advised to use Field-Weighted Citation Impact to avoid this drop, if it is of concern

Useful partner metrics are:

- **Field-Weighted Citation Impact** is a natural complement to Citations per Publication, taking into account behavioral differences between disciplines
- **Field-Weighted Citation Impact, Outputs in Top Percentiles, and Publications in Top Journal Percentiles** avoid display of the “dip” in recent years due to little time having passed to receive citations since publication
- **Cited Publications, that provides a measure of the reliability that an entity’s publications will subsequently be used to support other research**, but that is not affected by a high number of citations received by 1 or a few publications
- **Collaboration Impact and Academic-Corporate Collaboration Impact** also measure Citations per Publication, and are a natural complement. They focus on subsets of publications within a data set with particular collaboration characteristics
- **The set of “time-independent metrics”** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 3**.

Example 3: Citation Count, Cited Publications and Citations per Publication

Scenario: The user would like to calculate the Citation Count, Cited Publications, or Citations per Publication of an entity that consists of 6 publications, and has selected the following viewing and calculation options.

Selected Publication Year range	2005 to 2013
Selected Publication Types	Articles, reviews and editorials
Selected Research Area	Medicine

Entity with 6 Publications			
Publication Identity	Publication 1	Publication 2	
Publication Year	2008	2007	
Publication Type	Review	Editorial	
Total citations received by this publication	0	4	
Journal Sub-Category (or Sub-Categories)	Anthropology	Emergency Medicine	Management Science and Operations Research
Journal Main Category (or Categories)	Social Sciences	Medicine	Decision Sciences
Do publications match user-selected options?			
Step 1	Does the Journal Category of the publication match the selected Research Area?	No	Yes No
Step 2	Does this publication fall in the selected Publication Year range?	Yes	Yes
Step 3	Does the Publication Type match the selected Publication Types?	Yes	Yes
Step 4	Does the publication pass each of steps 1, 2 and 3?	No	Yes
Question: How do I calculate Scholarly Output?	Answer: Count the number of publications that received a "yes" in step 4.	Scholarly Output = 4	
Question: How do I calculate Citation Count?	Answer: Retrieve total citations received for the publications that received a "yes" in step 4.	N/A	4
	– Sum the citations received by those publications that received a "yes" in step 4.	Citation Count = 15	
Question: How do I calculate Cited Publications?	Answer: Have the publications that passed step 4 received at least 1 citation?	N/A	Yes
	– For "Absolute number" display option, sum the publications that have received at least 1 citation.	Cited Publications = 3	
	– For "Percentage" display option, divide "Absolute number" by Scholarly Output.	Cited Publications = 75.0%	
Question: How do I calculate Citations per Publication?	Answer: Divide "Citation Count" by "Scholarly Output"	Citations per Publication = 3.8	

Entity with 6 Publications				
Publication 3		Publication 4	Publication 5	Publication 6
2005		2004	2008	2010
Article		Article	Article	Review
7		9	0	4
Anatomy	Information Systems and Management	Immunology and Allergy	General Medicine	Emergency Medicine
Medicine	Decision Sciences	Medicine	Medicine	Medicine
Yes	No	Yes	Yes	Yes
Yes		No	Yes	Yes
Yes		Yes	Yes	Yes
Yes		No	Yes	Yes
7		N/A	0	4
Yes		N/A	No	Yes

4.8 Metric: Number of Citing Countries

Number of Citing Countries in SciVal indicates the geographical visibility of an entity's publications: from how many distinct countries have this entity's publications received citations?

Number of Citing Countries is a:

- Citation Impact metric
- “Power metric”

SciVal often displays Number of Citing Countries in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Compare smaller entities**, such as Groups of Researchers and Publication Sets, where differences in the number of citing countries are most likely to be evident
- Provide impressive figures to **showcase the performance of a relatively large entity**, when this metric is likely to give high numbers
- **Benchmark the geographical visibility of the publication portfolios of related entities**, such as
 - Collaboration networks in a given field of research
 - Researchers in a common field of research and with similar career lengths
 - Scenario models of a research institute, created to investigate the effect of recruiting different researchers
- **Provide evidence of extensive geographical appeal** of a body of work by indicating the diversity of the geographical sources of citations
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal
- Look at publishing activity in a way that is **difficult to manipulate**

This metric should be used with care when:

- **Comparing large entities**, such as Institutions and Groups of Countries, which will likely publish so many publications that receive so many citations, that the number of citing countries will approach the maximum and differences may not be visible. Users are advised to narrow their focus to a slice of a large entity when using this metric, such as by applying the Research Area filter
- **Benchmarking entities of obviously different sizes**, such as institutions and small departments, when this “Power metric” will most closely reflect entity size rather than differences in global visibility. It is not advised to use this metric to compare entities of different sizes

-
- **Benchmarking the collaboration of entities in different disciplines:**
 - The international appeal of publications from different disciplines may be inherently different, such as between national literature and chemistry, or local history and computer science
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities
 - **Understanding the magnitude of the number of citations received** by an entity's publications. Users are advised to use Citation Count or Citations per Publication to communicate information about magnitude
 - **Entities are small and there may be gaps in their output within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on apparent global visibility, whereas the effect of 1 or a few missing publication(s) from a large data set may be acceptable
 - The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher. Consider also limiting the use of Number of Citing Countries to comparing somewhat larger data sets in the same discipline where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison
 - **Investigating the performance of publications in the very early stages of a new strategy, or of early-career researchers**, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. It is advised to use Journal Count or other time-independent metrics in these situations

Useful partner metrics are:

- **Citation Count and Citations per Publication**, which communicate information about the magnitude of the number of citations received
- **Collaboration**, that shows the extent of international co-authorship of an entity's scholarly output and is a natural complement to Number of Citing Countries' view on geographical visibility
- **The set of all other "Power metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Citation Count, Cited Publications ("Absolute number"), Collaboration ("Absolute number"), Academic-Corporate Collaboration ("Absolute number"), Outputs in Top Percentiles ("Absolute number"), Publications in Top Journal Percentiles ("Absolute number"), and *h*-indices
- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Example 4: Number of Citing Countries

Scenario: The user would like to calculate the Number of Citing Countries of an entity that consists of 6 publications. They have not selected any viewing or calculation options. Say that this entity has received 6 citations from publications A, B, C, D, E and F.

	Entity with 6 Publications					
	Publication 1	Publication 2	Publication 3	Publication 4	Publication 5	Publication 6
Cited by: Publication A	Yes				Yes	
Cited by: Publication B			Yes		Yes	
Cited by: Publication C	Yes	Yes			Yes	Yes
Cited by: Publication D	Yes	Yes		Yes		Yes
Cited by: Publication E	Yes	Yes		Yes		Yes
Cited by: Publication F		Yes			Yes	

Scenario: The citing publications A, B, C, D, E and F have the following affiliation information:

Citing Publication	Authors	Institutions	Country
Publication A	A2	I4	C2
Publication B	A1	I1	C1
	A3	I4	C2
	A3	I2	C1
Publication C	A1	I1	C1
	A1	I3	C2
	A2	I2	C1
Publication D	A1	I1	C1
	A4	I1	C1
Publication E	A1	I1	C1
	A3	I5	C3
	A5	I1	C1
Publication F	A1	I1	C1
	A3	I8	C4
	A2	I1	C1

Question: How do I calculate the number of Citing Countries?

Answer: Count the number of distinct countries in the affiliations of the citing publications.

Number of Citing Countries = 4

4.9 Metric: Field-Weighted Citation Impact 🌐

Field-Weighted Citation Impact in SciVal indicates how the number of citations received by an entity's publications compares with the average number of citations received by all other similar publications in the data universe: how do the citations received by this entity's publications compare with the world average?

- A Field-Weighted Citation Impact of 1.00 indicates that the entity's publications have been cited exactly as would be expected based on the global average for similar publications; the Field-Weighted Citation Impact of "World", or the entire Scopus database, is 1.00
- A Field-Weighted Citation Impact of more than 1.00 indicates that the entity's publications have been cited more than would be expected based on the global average for similar publications; for example, 2.11 means 111% more cited than world average
- A Field-Weighted Citation Impact of less than 1.00 indicates that the entity's publications have been cited less than would be expected based on the global average for similar publications; for example, 0.87 means 13% less cited than world average

Similar publications are those publications in the Scopus database that have the same publication year, publication type, and discipline, as represented by the Scopus journal classification system:

- Publications can be assigned to a classification system in 2 ways:
 - **"Journal-driven" assignment** assumes that every publication within a journal fits within the same discipline(s) as the journal's scope. Each publication automatically adopts the subject classifications that are assigned to the journal. This method of assignment is suitable for journals that are focused in a core field, and do not tend to include publications that are also relevant to other fields
 - **"Publication-driven" assignment** assumes that publications within a journal may have additional or different relevance to fields outside the core focus of the journal's scope. Publication-driven assignment offers the benefit of being able to assign individual publications from a journal separately to their relevant classifications. This is important for publications in multi-disciplinary journals
- Field-Weighted Citation Impact uses "publication-driven" assignment
- Publications are allocated to the classification Sub-category level, and can be allocated to more than 1 Sub-category. When we calculate the expected citations for similar publications, it is important that these multi-category publications do not exert too much weight; for example, if a publication P belongs to both in both parasitology and microbiology, it should not have double the influence of a publication that belongs to only one or the other Sub-category. This is accounted for in SciVal by distributing publication and citation counts equally across multiple journal categories; publication P would be counted as 0.5 publications for each of parasitology and microbiology, and its citations would be shared equally between these Sub-categories

Field-Weighted Citation Impact is a:

- Citation Impact metric
- Snowball Metric

SciVal often displays Field-Weighted Citation Impact in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received. The citations received in the year in which an item was published, and the following 3 years, are counted for this metric.

This metric is useful to:

- **Benchmark entities regardless of differences in their size, disciplinary profile, age, and publication-type composition**, such as
 - An institution and departments (Groups of Researchers) within that institution
 - A country and small research institutes within that country
 - A geographical region and countries within that region
- **Easily understand the prestige of an entity's citation performance** by observing the extent to which its Field-Weighted Citation Impact is above or below the world average of 1.00
- Present citation data in a way that **inherently takes into account the lower number of citations received by relatively recent publications**, thus avoiding the dip in recent years seen with Citation Count and Citations per Publication
- **Gain insight into the relative citation performance of an entity in a discipline with relatively poor Scopus coverage**, since gaps in the database will apply equally to the entity's publications and to the set of similar publications
- Use as a **default to view citation data**, since it takes multiple variables that can affect other metrics into account
- Look at publishing activity in a way that is **difficult to manipulate**

This metric should be used with care when:

- **Information about the magnitude of the number of citations received** by an entity's publications is important. In these situations, it is advised to use Citation Count or Citations per Publication
- **Demonstrating excellent performance to those who prefer to see high numbers**; Citation Count or Citations per Publication would be more suitable in these circumstances
- **Entities are small so that the metric may fluctuate significantly and appear unstable** over time, even when there is complete Scopus coverage. Field-Weighted Citation Impact calculates an average value, and these types of calculations are strongly influenced by outlying publications in a small data set

-
- **Trust needs to be built** in the metrics in SciVal. This calculation accounts for multiple normalizations, and the generation of the average citations for similar publications requires a view on the entire Scopus database which will be difficult for a user to validate. Users are advised to select simpler metrics, such as Citation Count or Citations per Publication, if trust in the accuracy of the SciVal calculations needs to be built
 - **Uncovering the performance of publications in the very early stages of a new strategy, or of early-career researchers**, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. It is advised to use on of the time-independent metrics, such as Academic-Corporate Collaboration, in these cases
 - **Completely answering every question about performance.** Field-Weighted Citation Impact is a very useful metric, but using it to the exclusion of other metrics severely restricts the richness and reliability of information that a user can draw from SciVal

Useful partner metrics are:

- **Citation Count and Citations per Publication.** They indicate the magnitude of the number of citations received, to complement this relative view offered by Field-Weighted Citation Impact. They are also simple and offer transparency on the underlying data to build trust in SciVal's metrics calculations
- **The set of “time-independent metrics”** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Mathematical notation:

The Field-Weighted Citation Impact (FWCI) for a set of N publications is defined as:

$$FWCI \equiv \frac{1}{N} \sum_{i=1}^N \frac{c_i}{e_i}$$

c_i = citations received by publication i

e_i = expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than 1 discipline, the harmonic mean is used to calculate e_i . For a publication i that is part of 2 disciplines:

$$\frac{1}{e_i} \equiv \frac{1}{2} \left(\frac{1}{e_A} + \frac{1}{e_B} \right)$$

e_A , e_B = fractional counts of publications and citations, so that publication i will be counted as 0.5 publications in each of e_A and e_B , and the citations it has received will also be shared between **A** and **B**.

Example 5: Field-Weighted Citation Impact

Scenario: The user would like to calculate the Field-Weighted Citation Impact of an entity that consists of 3 publications. They have not selected any viewing or calculation options.

Entity with 3 Publications				
Publication Identity	Publication 1	Publication 2		Publication 3
Publication Year (pub year)	2009	2010		2013
Publication Type	Article	Review		Erratum
Journal category(ies)	Immunology	Immunology	Parasitology	Parasitology
Step 1	– Compute number of citations received by publications in the entity.	2	12	0
	– Actual citations received in pub year	3	23	N/A (example prepared in 2013)
	– Actual citations received in 1st year after pub year	13	28	N/A (example prepared in 2013)
	– Actual citations received in 2nd year after pub year	23	45	N/A (example prepared in 2013)
	– Actual citations received in 3rd year after pub year	2+3+13+23	12+23+28+45	
	– Actual citations received by the individual publication in pub year plus following 3 years	41	108	0
Step 2	– Compute expected number of citations received by similar publications.	7,829.6	1,349.8	161.9
	– Number of publications in database published in same year, of same type, and within same discipline as Publication 1, 2, or 3	141,665.2	35,770.8	2,161.5
	– Total citations received in pub year plus 3 years by all publications in database published in same year, of same type, and within same journal category(ies) as Publication 1, 2, or 3	141,665.2 / 7,829.6	35,770.8 / 1,349.8	2,161.5 / 161.9
	– Average citations per publication for all publications in database published in same year, of same type, and within same subject category as Publication 1, 2, or 3	18.09	26.50	13.35
Step 3	– Use harmonic mean to compute expected number of citations for publications covered in multiple categories.	$2/(1/26.5 + 1/13.4)$		
	– Combined average citations per publication for publications indexed in multiple journal categories	17.76		
Step 4	Compute ratio of actual (result of step 1) to expected (result of step 2 or 3) citations for each of Publications 1, 2 and 3.	41/18.1 2.27	108/17.76 6.08	No citations received or expected => 0 0.00
Step 5	Take arithmetic mean of the results of step 4 to calculate Field-Weighted Citation Impact for this entity	Arithmetic Mean $(2.27 + 6.08 + 0.00) / 3$ Field-Weighted Citation Impact = 2.78		

0.9 of an article? 0.3 of a review?

Publications in Scopus can belong to more than one journal category. When calculating the expected citations per publication as input to Field-Weighted Citation Impact, we count multi-category publications partially within each of their journal categories, and distribute their citations accordingly. No weighting between journal categories is applied; publications are assumed to belong equally to each category they are indexed in.

For example, if a publication with 3 citations belongs to both Parasitology and Microbiology journal categories, it is counted as 0.5 publications with 1.5 citations in Parasitology, and as 0.5 publications with 1.5 citations in Microbiology.

This fractional distribution ensures that all publications have equal weight, regardless of the number of journal categories.

The net effect of adding up these fractional publication and citation counts is 0.9 of an article, and 0.3 of a review, as seen here in the statistics on the SciVal data set.

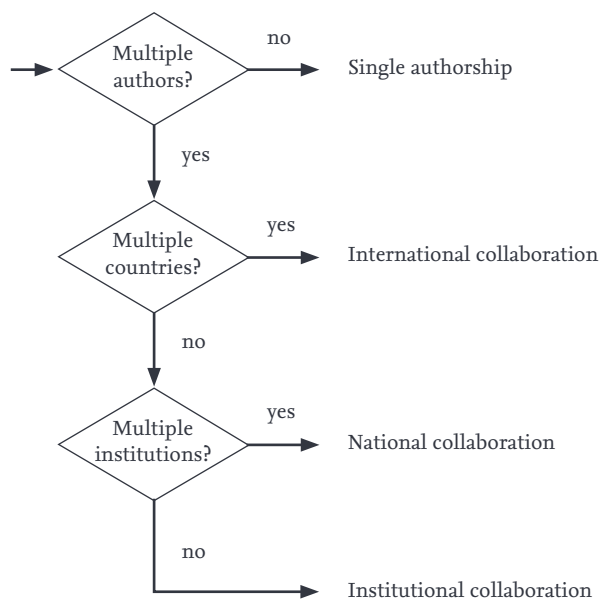
Statistics on SciVal data set			
Pub Year	Number of articles	Citations received in pub year +3 subsequent years	Citations per article
Immunology, Articles			
2008	11,760.0	266,081.8	22.63
2009	7,829.6	141,665.2	18.09
2010	7,729.5	103,824.8	13.43
2011	8,355.9	73,485.9	8.79
2012	8,600.2	31,669.5	3.68
2013	6,552.2	4,253.4	0.65
Immunology, Reviews			
2008	1,586.3	71,346.9	44.98
2009	1,270.0	44,530.6	35.06
2010	1,349.8	35,770.8	26.50
2011	1,305.3	20,765.2	15.91
2012	1,497.8	9,606.1	6.41
2013	1,067.0	1,056.1	0.99
Parasitology, Reviews			
2008	227.3	4,301.6	18.92
2009	185.9	3,484.2	18.74
2010	161.9	2,161.5	13.35
2011	195.3	1,413.2	7.24
2012	205.1	700.1	3.41
2013	153.0	104.1	0.68
Parasitology, Errata			
2008	24.7	1.2	0.05
2009	15.9	5.2	0.32
2010	13.6	1.0	0.07
2011	21.2	7.5	0.35
2012	15.5	0.3	0.02
2013	8.3	0.0	0.00
(Scopus data as of 11 November, 2013)			

4.10 Metric: Collaboration

Collaboration in SciVal indicates the extent to which an entity's publications have international, national, or institutional co-authorship, and single authorship.

Each publication is assigned to 1 of 4 collaboration types, based on its affiliation information: international, national, institutional, or single authorship. A single publication may of course display each of international, national and institutional collaboration in its affiliation information, but a single collaboration type is assigned to ensure that the sum of an entity's publications across the 4 categories adds up to 100.0% of the publications with the necessary affiliation information.

The assignment of geographical collaboration type is performed using the following cascading decision tree:



Collaboration is a:

- Collaboration metric
- Snowball Metric
- “Power metric” when the “Absolute number” option is selected, but not when the “Percentage” option is selected

This metric is useful to:

- Explore the **extent of international and other types of collaboration** within a data set
- **Benchmark the collaboration of entities of different sizes**, but in similar disciplines
 - It is advised to select the “Percentage” option when comparing entities of different sizes, to normalize for this variable
- **Showcase extensive international collaboration** that may underpin a set of scholarly output
- Look at publishing activity in a way that is **difficult to manipulate**
- **Investigate collaborative activity very early in a new strategy, or for an early-career researcher**, since the affiliation data underlying this metric do not require time to accumulate to reliable levels in the same way as citation data do
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

This metric should be used with care when:

- **Benchmarking the collaboration of entities in different disciplines:**
 - The typical collaborative behavior of academics may differ between disciplines, such as mathematics and medicine, or humanities and molecular biology
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution, or a Group of Institutions such as a US state, it is advised to apply the Research Area filter to focus on one field that is common between all the entities

-
- **Entities are small and there may be gaps in their output within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on apparent partnerships, whereas the effect of 1 or a few missing publication(s) from a large data set may be acceptable
 - The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher, or to limit the use of Collaboration to comparing larger data sets where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison
 - **Understanding activity in a discipline with an obvious national focus**, such as Finnish literature or cultural studies, when institutional and national collaboration may be of more use than international collaboration

Useful partner metrics are:

- **Number of Citing Countries**, which indicates the diversity of countries in which Researchers have built on an entity's publications. It is likely to be higher if an entity's international collaboration activity involves several countries, rather than just 1 or 2, even though both situations can give 100.0% international collaboration
- **Collaboration Impact**, which calculates the average Citations per Publication for publications with different types of geographical collaboration, and indicates how beneficial these collaborations are with respect to citation impact
- **The set of all other "Power metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Citation Count, Cited Publications ("Absolute number"), Number of Citing Countries, Academic-Corporate Collaboration ("Absolute number"), Outputs in Top Percentiles ("Absolute number"), Publications in Top Journal Percentiles ("Absolute number"), and *h*-indices
- **The set of all other "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 6**.

4.11 Metric: Collaboration Impact

Collaboration Impact in SciVal indicates the citation impact of an entity's publications with particular types of geographical collaboration: how many citations do this entity's internationally, nationally, or institutionally co-authored publications receive, as well as those with a single author?

Publications are assigned to 1 of 4 geographical collaboration types, as explained for Collaboration. This assignment applies to the entity's publications only, and the count of citations received is independent of the geographical collaboration status of the citing publications themselves; if an internationally collaborative publication is cited by another publication with single authorship, that citation is still counted.

This metric calculates the Citations per Publication for each type of geographical collaboration.

Collaboration Impact is a:

- Citation Impact metric
- Collaboration metric

SciVal often displays Collaboration Impact in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Benchmark the average influence of an entity's publications with particular types of geographical collaboration**, such as:
 - The citation impact of an entity's internationally collaborative publications with that of its institutionally collaborative output
 - The citation impact of internationally collaborative publications compared to that of non-collaborative publications with a single author
- **Compare the collaborative citation impact of entities of different sizes**, but from related disciplines
- **Demonstrate any benefit from establishing and maintaining collaborations** to the citation impact of an entity's Scholarly Output
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

This metric should be used with care when:

- **Comparing entities in different disciplines** where the citing behavior of academics may differ:
 - Citation counts tend to be higher in disciplines such as virology, whose academics tend to publish frequently and include long reference lists, than in law, for example; these differences reflect the differences in the behavior of researchers in distinct subject fields, and not differences in performance
 - It is not advisable to use this metric to compare entities in distinct disciplines without taking these differences into account
 - When comparing entities made up of a mixture of disciplines, such as an Institution or Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities
- **Understanding the citation impact of collaborative papers of small entities**, when there may be gaps in their output within the Scopus coverage:
 - A single missing publication from a small data set may significantly distort the apparent performance, whereas the buffering effect of a larger data set may compensate for 1 or 2 missing publications
 - The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher. It is also advisable to limit the use of these metrics to comparing larger data sets where gaps in the database coverage likely have a similar effect on all entities being viewed and will not invalidate the comparison
- **Entities are small so that the metric may fluctuate significantly and appear unstable** over time, even when there is complete Scopus coverage. Collaboration Impact calculates an average value, and these types of calculations are strongly influenced by outlying publications in a small data set
- **Understanding the performance of publications in the very early stages of a new strategy, or of early-career researchers**, where the short time that has passed since publication will reduce the reliability of citation information as an input into decisions. These situations can be addressed by metrics that are useful immediately upon publication, such as Collaboration or Publications in Top Journal Percentiles

Useful partner metrics are:

- **Citations per Publication**, that calculates the average citation impact across all publications within an entity regardless of their geographical collaboration status. This metric could be a useful benchmark against which to judge the extent of benefit of different types of geographical collaboration
- **Collaboration**, which indicates the extent to which an entity's publications have international, national, or institutional co-authorship, and single authorship. This is a natural partner metric to Collaboration Impact
- **The set of “time-independent metrics”** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 6**.

4.12 Metric: Academic-Corporate Collaboration

Academic-Corporate Collaboration in SciVal indicates the degree of collaboration between academic and corporate affiliations: to what extent are this entity's publications co-authored across the academic and corporate, or industrial, sectors?

A publication either exhibits academic-corporate collaboration, or it does not. This assignment is made based on the organization-type with which Scopus tags each affiliation.

This metric calculates the Citations per Publication for collaborative and non-collaborative papers. Academic-Corporate Collaboration is a:

- Collaboration metric
- “Power metric” when the “Absolute number” option is selected, but not when the “Percentage” option is selected

This metric is useful to:

- Investigate the **degree of collaboration between the academic and corporate sectors** within a data set
- **Benchmark the cross-sector collaboration of entities of different sizes**, but in related disciplines, such as large and small research teams, or large and small Centers of Excellence
 - It is advised to select the “Percentage” option when comparing entities of different sizes, to normalize for this variable
- **Showcase extensive collaboration between academia and industry** that may underpin a set of Scholarly Output
- Look at publishing activity in a way that is **difficult to manipulate**
- **Investigate collaborative activity very early in a new strategy, or for an early-career researcher**, since the affiliation data underlying this metric do not require time to accumulate to reliable levels in the same way as citation data do
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

This metric should be used with care when:

- **Benchmarking the extent of academic-corporate collaboration of entities in different disciplines:**
 - The opportunity or desire to collaborate outside the sector may differ, such as between econometrics and drug discovery, or the philosophy of science and toxicology
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities
- **Entities are small and there may be gaps in their output within the Scopus coverage:**
 - A single missing publication from a small data set may have a significant negative impact on apparent cross-sector partnerships, whereas the effect of 1 or a few missing publication(s) from a large data set may be acceptable
 - The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher, or to limit the use of Academic-Corporate Collaboration to comparing larger data sets where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison
- Investigating activity in a discipline with **focus outside the interest of industry**, such as history. It is not advised to use this metric in such a situation

Useful partner metrics are:

- **Academic-Corporate Collaboration Impact**, which calculates the Citations per Publication for publications with and without academic-corporate collaboration, and indicates how beneficial this cross-sector collaboration is with respect to citation impact
- **The set of all other “Power metrics”** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Citation Count, Cited Publications (“Absolute number”), Number of Citing Countries, Collaboration (“Absolute number”), Outputs in Top Percentiles (“Absolute number”), Publications in Top Journal Percentiles (“Absolute number”), and *h*-indices
- **The set of all other “time-independent metrics”** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, and Publications in Top Journal Percentiles

See **Example 6**.

4.13 Metric: Academic-Corporate Collaboration Impact

Academic-Corporate Collaboration Impact in SciVal indicates the citation impact of an entity's publications with or without both academic and corporate affiliations: how many citations do this entity's publications receive when they list both academic and corporate affiliations, versus when they do not?

A publication either exhibits academic-corporate collaboration, or it does not. This assignment applies to the entity's publications only, and the count of citations received is independent of the collaboration status of the citing publications themselves; if a publication that resulted from academic-corporate collaboration is cited by another publication with only academic affiliations, that citation is still counted.

Academic-Corporate Collaboration Impact is a:

- Citation Impact metric
- Collaboration metric

SciVal often displays Academic-Corporate Collaboration Impact in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Benchmark the average influence of an entity's publications with and without academic-corporate collaboration**, such as:
 - The citation impact of publications of academic research institutes that were published with industrial collaboration, compared to those that were not
 - The citation impact of publications of Researchers located in a corporate affiliation that were published with academic collaboration, compared to those that were industry-only
- **Compare the citation impact of these cross-sector publications between entities of different sizes**, but from related disciplines, such as large and small international networks of Researchers in neuroscience
- **Demonstrate any benefit from establishing and maintaining academic-corporate collaborations** to the citation impact of an entity's scholarly output
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

This metric should be used with care when:

- **Comparing entities in different disciplines** where the citing behavior of academics may differ:
 - Citation counts tend to be higher in disciplines such as cardiology, whose academics tend to publish frequently and include long reference lists, than in anthropology, for example; this reflects differences in the behavior of researchers in distinct subject fields, and not differences in performance
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution, a Country or a Group of Countries, it is advised to apply the Research Area filter to focus on one field that is common between all the entities
- **Understanding the citation impact of collaborative papers of small entities**, when there may be gaps in their output within the Scopus coverage:
 - A single missing publication from a small data set may significantly distort the apparent performance, whereas the buffering effect of a larger data set may compensate for 1 or 2 missing publications
 - The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher. It is also advisable to limit the use of these metrics to comparing larger data sets where gaps in the database coverage likely have a similar effect on all entities being viewed and will not invalidate the comparison
- **Entities are small so that the metric may fluctuate significantly and appear unstable** over time, even when there is complete Scopus coverage. Academic-Corporate Collaboration Impact calculates an average value, and these types of calculations are strongly influenced by outlying publications in a small data set
- **Understanding the performance of publications in the very early stages of a new strategy, or of early-career researchers**, where the short time that has passed since publication will reduce the reliability of citation information as an input into decisions. These situations can be addressed by metrics that are useful immediately upon publication, such as Academic-Corporate Collaboration or Publications in Top Journal Percentiles

Useful partner metrics are:

- **Citations per Publication**, that calculates the average citation impact across all publications within an entity regardless of their academic-corporate collaboration status. This metric could be a useful benchmark against which to judge the extent of benefit of different types of collaboration
- **Academic-Corporate Collaboration**, which indicates the degree of collaboration between academic and corporate affiliations. This is a natural partner metric to Academic-Corporate Collaboration Impact
- **The set of “time-independent metrics”** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See **Example 6**.

Example 6: Collaboration, Collaboration Impact, Academic-Corporate Collaboration, and Academic-Corporate Collaboration Impact

Scenario: The user would like to calculate the Collaboration, Collaboration Impact, Academic-Corporate Collaboration, or Academic-Corporate Collaboration Impact of an entity that consists of 6 publications.

Selected Publication Year range	2005 to 2012
Selected Publication Types	Articles, reviews and editorials
Selected Research Area	Medicine

Entity with 6 Publications					
Publication Identity	Publication 1		Publication 2		
Publication Year	2011		2009		
Total citations received by this publication	8		11		
Publication Type	Letter		Review		
Journal Sub-Category	Anatomy		Dermatology		
Journal Main Category	Medicine		Medicine		
Authors	Author 1	Author 2	Author 1	Author 3	
Institutions	Institution 1	Institution 1	Institution 1	Institution 4	Institution 2
Countries	Country 1	Country 1	Country 1	Country 2	Country 1
Is the institution academic?	Yes	Yes	Yes	No	No
Is the institution corporate?	No	No	No	No	Yes
Step 1	Does the Journal Category of the publication match the selected Research Area?		Yes		
Step 2	Does this publication fall into the selected Publication Year range?		Yes		
Step 3	Does the Publication Type match the selected Publication Types?		Yes		
Step 4	Does the publication pass each of steps 1, 2 and 3?		Yes		
Question: How do I calculate Scholarly Output?	Answer: Count the number of publications that received a "yes" in step 4.		Scholarly Output = 4		
Question: How do I calculate Collaboration?	Answer: Assign collaboration type to each publication that received a "yes" in step 4 using the decision tree. (i) Does the publication have a single author? If Yes, assign as single Authorship. (ii) Do the unassigned publications have more than one country in the affiliation information? If yes, assign as International Collaboration. (iii) Do the unassigned publications have more than one institution in the affiliation information? If yes, assign as National Collaboration. (iv) Assign all remaining publications as Institutional Collaboration.		International Collaboration		
	For "Absolute number" display option, sum the publications within each group.		Single Authorship = 1 International Collaboration = 2 National Collaboration = 0 Institutional Collaboration = 1		
	For "Percentage" display option, divide "Absolute Number" within each group by Scholarly Output of the entity.		Single Authorship Impact = $(1/4) \times 100 = 25.0\%$ International Collaboration Impact = $(2/4) \times 100 = 50.0\%$ National Collaboration Impact = $(0/4) \times 100 = 0.0\%$ Institutional Collaboration Impact = $(1/4) \times 100 = 25.0\%$		
Question: How do I calculate Collaboration Impact?	Answer: Divide total citations received by the publications by absolute number of each type of collaboration.		Single Authorship Impact = $1/1 = 1.0$ International Collaboration Impact = $(11+6)/2 = 8.5$ National Collaboration Impact = $0/0 = 0.0$ Institutional Collaboration Impact = $4/1 = 4.0$		
Question: How do I calculate Academic-Corporate Collaboration?	Answer: Do the publications that received a "yes" in step 4 have both an academic and a corporate institution in the affiliation information?		Yes		
	For "Absolute Number" display option, sum the publications with and without academic-corporate collaboration.		With collaboration = 1 Without collaboration = 3		
	For "Percentage" display option, divide "Absolute Number" within each group by Scholarly Output of the entity.		With collaboration = $(1/4) \times 100 = 25.0\%$ Without collaboration = $(3/4) \times 100 = 75.0\%$		
Question: How do I calculate Academic-Corporate Collaboration Impact?	Answer: Divide total citations received by the publications by absolute number of each type of collaboration.		Impact with collaboration = $11/1 = 11.0$ Impact without collaboration = $(1+4+6) / 3 = 3.7$		

Entity with 6 Publications										
Publication 3			Publication 4		Publication 5			Publication 6		
2011			2011		2010			2003		
1			4		6			12		
Article			Editorial		Review			Article		
Anatomy			Internal Medicine		General Medicine			Emergency Medicine		
Medicine			Medicine		Medicine			Medicine		
Author 1			Author 1	Author 4	Author 1		Author 4	Author 1		Author 2
Institution 1	Institution 3	Institution 5	Institution 1	Institution 1	Institution 1	Institution 5	Institution 1	Institution 1	Institution 2	Institution 1
Country 1	Country 2	Country 3	Country 1	Country 1	Country 1	Country 3	Country 1	Country 1	Country 1	Country 1
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
No	No	No	No	No	No	No	No	No	Yes	No
Yes			Yes		Yes			Yes		
Yes			Yes		Yes			No		
Yes			Yes		Yes			Yes		
Yes			Yes		Yes			No		
Single Authorship			Institutional Collaboration		International Collaboration					
No			No		No					

4.14 Metric: Outputs in Top Percentiles

Outputs in Top Percentiles in SciVal indicates the extent to which an entity's publications are present in the most-cited percentiles of a data universe: how many publications are in the top 1%, 5%, 10% or 25% of the most-cited publications?

The entire Scopus database, or "World", is the default data universe used to generate this metric:

- The citation counts that represent the thresholds of the 1%, 5%, 10% and 25% most-cited papers in Scopus per Publication Year are calculated. Sometimes the same number of citations received by the publication at, say, the 10% boundary has been received by more than 10% of publications; in this case, all of the publications that have received this number of citations are counted within the top 10% of the Scopus data universe, even though that represents more than 10% by volume
- SciVal uses these citation thresholds to calculate the number of an entity's publications that fall within each percentile range

SciVal users have the option to select from additional data universes within which to generate this metric, such as countries. If a user selects China, or Canada, or the United Kingdom, then SciVal will use the citation count thresholds at the boundaries of the 1%, 5%, 10% and 25% most-cited papers in China, or Canada, or the United Kingdom per Publication Year in this metric calculation. This selection does not filter the publications of the entity upon which the calculation is performed, and it affects only the citation thresholds.

Use of the Publication Type filter affects the publications in the data universe that are used to generate the citation thresholds, as well as the publications of the entity upon which the calculation is performed. The exclusion of self-citations affects only the entity, and not the data universe.

Outputs in Top Percentiles can only be calculated for the current year from the first data snapshot on or after 1 July. It will be displayed as a null value until this date is reached. This metric depends on being able to divide the publications into 100 percentiles, and this level of division is not possible earlier in the publication year when items just published have received very few citations.

Outputs in Top Percentiles is a:

- Citation Impact metric
- Snowball Metric
- “Power metric” when the “Absolute number” option is selected, but not when the “Percentage” option is selected

SciVal often displays Outputs in Top Percentiles in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Benchmark the contributions towards the most influential, highly cited publications in the world** of entities of different sizes, but in similar disciplines
 - It is advised to select the “Percentage” option when comparing entities of different sizes, to normalize for this variable
- **Distinguish between entities whose performance seems similar** when viewed by other metrics, such as Scholarly Output, Citations per Publication, or Collaboration
- **Showcase the performance of a prestigious entity** whose publications are amongst the most cited and highly visible publications of the scholarly world
- Present citation data in a way that **inherently takes into account the lower number of citations received by relatively recent publications**, thus avoiding the dip in recent years seen with Citation Count and Citations per Publication

This metric should be used with care when:

- **Comparing entities in different disciplines:**
 - Citation counts tend to be higher in disciplines such as immunology and microbiology, whose academics tend to publish frequently and include long reference lists, than in mathematics, where publishing 1 item every 5 years that refers to 1 or 2 other publications is common; these differences reflect the distinct behavior of researchers in distinct subject fields, and not differences in performance
 - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences
 - When comparing entities made up of a mixture of disciplines, such as an Institution or Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities
- **Entities are small and there may be gaps in their output within the Scopus coverage.** A single missing highly cited publication from a small data set will have a significant negative impact on apparent performance. Although it is relatively unlikely that such prominent publications are not indexed by Scopus, we advise users to be vigilant and to bear this possible limitation in mind
- There is a **concern that excessive self-citations may be artificially inflating the number of publications that appear in the top percentiles.** Users can judge whether the level of self-citations is higher than normal by deselecting the “Include self-citations” option
- Understanding the **status of publications of an early-career researcher, or those resulting from the first stages of a new strategy**, where insufficient time may have passed to ensure that presence in top citation percentiles is a reliable indicator of performance. These situations can be addressed by metrics that are useful immediately upon publication, such as Publications in Top Journal Percentiles
- **Trust needs to be built** in the metrics in SciVal. The citation thresholds may depend on the entire Scopus database, and it will be difficult for a user to validate these boundaries. Users are advised to select simpler metrics, such as Citations per Publication, if trust in the accuracy of the SciVal calculations needs to be built

Useful partner metrics are:

- **Cited Publications**, which indicates the reliability with which an entity's output is built on by subsequent research by counting publications that have received at least 1 citation. It is not affected by 1 or a few very highly cited publications
- **Publications in Top Journal Percentiles**, which indicates the extent to which an entity's publications are present in the most-cited journals in the data universe, and is independent of the citations received by the publications themselves. This is a natural partner metric
- **The set of all other "Power metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Citation Count, Cited Publications ("Absolute number"), Number of Citing Countries, Collaboration ("Absolute number"), Academic-Corporate Collaboration ("Absolute number"), Publications in Top Journal Percentiles ("Absolute number"), and *h*-indices
- **The set of "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Example 7: Outputs In Top Percentiles

Scenario: The user would like to calculate the Outputs in the Top Percentiles of an entity that consists of 6 publications, and has selected the following viewing and calculation options:

Selected Publication Year range	2004 to 2013
Selected Publication Types	Articles, reviews and editorials
Selected Research Area	Chemistry
Selected Percentile Level	10%
Selected data universe	World

Entity with 6 Publications		
Publication Identity	Publication 1	Publication 2
Publication Year	2011	2009
Total citations received by this publication	10	42
Publication Type	Article	Review
Journal Sub-Category	Organic Chemistry	Pharmaceutical Science
Journal Main Category	Chemistry	Pharmacology, Toxicology & Pharmaceuticals

Do publications match user-selected options?

Step 1	Does the Journal Category of the publication match the selected Research Area?	Yes	No
Step 2	Does this publication fall in the selected Publication Year range?	Yes	Yes
Step 3	Does the Publication Type match the selected Publication Types?	Yes	Yes
Step 4	Does the publication pass each of steps 1, 2 and 3?	Yes	No

Question: How do I calculate Scholarly Output?

Answer: Count the number of publications that received a “yes” in step 4.

Scholarly Output = 4

Question: How do I calculate Outputs in Top Percentiles?

Answer: Look up the 10% citation threshold for this Publication Year of publications that received a “yes” in step 4.

8

N/A

– Was the publication cited at least as many times as the threshold?

Yes

N/A

– For “Absolute number” display option, count the publications that received a “yes” in the previous step.

Outputs in Top Percentiles = 3

– For “Percentage” display option, divide “Absolute Number” by the Scholarly Output of the entity.

Outputs in Top Percentiles = (3/4)*100 = 75.0%

This table shows the number of times a publication must be cited to be in the top 10% for its Publication Year.

10% World Citation Thresholds					
Year	Citations	Year	Citations	Year	Citations
1996	44	2003	42	2010	13
1997	45	2004	39	2011	8
1998	47	2005	33	2012	4
1999	49	2006	30	2013	1
2000	49	2007	26	2014	0
2001	45	2008	22		
2002	44	2009	18		

Entity with 6 Publications			
Publication 3	Publication 4	Publication 5	Publication 6
2011	2004	2010	2004
15	25	44	44
Editorial	Editorial	Review	Article
Geochemistry and Petrology	Spectroscopy	Inorganic Chemistry	Organic Chemistry
Earth & Planetary Sciences	Chemistry	Chemistry	Chemistry
No	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
No	Yes	Yes	Yes
N/A	39	13	39
N/A	No	Yes	Yes

4.15 Metric: Publications in Top Journal Percentiles

Publications in Top Journal Percentiles in SciVal indicates the extent to which an entity's publications are present in the most-cited journals in the data universe: how many publications are in the top 1%, 5%, 10% or 25% of the most-cited journals indexed by Scopus?

The most-cited journals are defined by the journal metrics SNIP (Source-Normalized Impact per Paper) or SJR (SCImago Journal Rank). This means that the data universe is the set of items indexed by Scopus that have a journal metric and so can be organized into percentiles; this excludes publications in stand-alone books and trade publications, which do not have journal metrics.

All items indexed by Scopus that have a SNIP or SJR value form the data universe used to generate this metric:

- The SNIP or SJR values at the thresholds of the top 1%, 5%, 10% and 25% most-cited journals in Scopus per Publication Year are calculated. It is possible that the same journal metric value received by the indexed item at, say, the 10% boundary has been received by more than 10% of indexed items; in this case, all items with this journal metric value are counted within the top 10% of the data universe, even though that represents more than 10% items by volume. This is less likely to happen than for the Outputs in Top Percentiles thresholds, because the journal metrics are generated to 3 decimal places which reduces the chance of items having the same value
- These thresholds are calculated separately for each journal metric, not once for a combination of both journal metrics
- SciVal uses these journal metric thresholds to calculate the number of an entity's publications within indexed items that fall within each percentile range
- Indexed items have multiple journal metric values, for distinct years. Which one is used in assigning publications to journal metrics thresholds?
 - The journal metric value matching the publication year of an item of scholarly output is used as far as possible
 - The first journal metric values are available for 1999. For scholarly output published in the range 1996-1999, the journal metric value for 1999 is used
 - Current year journal metric values are published during the course of the following year. For recent items whose journal metric values have not yet been published, the previous year's journal metric values are used until the current year's become available
- A publication may be counted in the top journal percentiles without itself ever having been cited. The citations received by an individual publication are irrelevant for this metric, which is based only on citations received by a journal or conference proceedings

SNIP and SJR are both field-normalized journal metrics, meaning that this metric can be used to compare the presence of publications in exceptional journals of entities working in different disciplines.

Publications in Top Journal Percentiles is a:

- Citation Impact metric
- “Power metric” when the “Absolute number” option is selected, but not when the “Percentage” option is selected

SciVal often displays Publications in Top Journal Percentiles in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- **Benchmark entities even if they have different sizes and disciplinary profiles:**
 - It is advised to select the “Percentage” option when comparing entities of different sizes, to normalize for this variable
 - SNIP and SJR are field-normalized metrics. Selecting one of these journal metrics will inherently account for differences in the behavior of academics between fields
- **Showcase the presence of publications in journals** that are likely to be perceived as the most prestigious in the world
- **Incorporate peer review into a metric**, since it is the judgment of experts in the field that determines whether a publication is accepted by a particular journal or not
- **Avoid the “dip” in recent years** seen with metrics like Citation Count and Citations per Publication
- Look at publishing activity in a way that is **difficult to manipulate**, since it relies on peer review
- **Investigate performance very early in a new strategy, or for an early-career researcher**, since journal data underpin this metric, and not the citations received by individual publications themselves which would require time to accumulate
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal. The journal metrics for the entire database are available for download in a spreadsheet from www.journalmetrics.com, so that a user can determine the thresholds themselves if required

This metric should be used with care when:

- **The objective is to judge an entity's publications based on their actual performance**, rather than based on the average of a journal:
 - A publication may appear in a journal with a very high SNIP or SJR value, and not itself receive any citations; even the most prestigious journals in the world contain publications that have never been cited
 - A publication may be very influential and have received many citations, without being part of a journal with a high SNIP or SJR value; journals which do not rank very highly in the data universe may still contain very highly cited publications
- **Benchmarking the performance of a Researcher who is the editor of a journal in the top percentiles**, since they can publish multiple editorials which all fall into these top percentiles. In this situation, it is advised to use the Publication Type filter to exclude editorials and ensure that the types of publications in the data sets being compared are consistent
- **Entities are small and there may be gaps in their output within the Scopus coverage**. A single missing publication from a small data set will have a significant negative impact on apparent performance. Although it is relatively unlikely that publications in such prominent journals are not indexed by Scopus, we advise users to be vigilant and to bear this possible limitation in mind
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

Useful partner metrics are:

- **Citation Count, Citations per Publication, Field-Weighted Citation Impact, and Outputs in Top Percentiles** which rely on the citations received by an entity's publications themselves, and not on the average performance of the journal
- **Outputs in Top Percentiles**, which indicates the extent to which an entity's publications are present in the most-cited percentiles of the data universe, but depends on the citations received by the publications themselves. This is a natural partner metric
- **The set of all other "Power metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Citation Count, Cited Publications ("Absolute number"), Number of Citing Countries, Collaboration ("Absolute number"), Academic-Corporate Collaboration ("Absolute number"), Outputs in Top Percentiles ("Absolute number"), and *h*-indices
- **The set of all other "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, and Academic-Corporate Collaboration

Example 8: Publication In Top Journal Percentiles

Scenario: The user would like to calculate the Publications in Top Journal Percentiles of an entity that consists of 6 publications, and has selected the following viewing and calculation options.

Selected Journal Metric:	SNIP
Selected Percentile Level:	10%

Entity with 6 Publications		
Publication Identity	Publication 1	Publication 2
Publication Year	2005	2006
Journal in which publication is published	<i>Perception</i>	<i>Journal of the American Medical Association</i>
Step 1	Retrieve journal's SNIP value for the Publication Year. 0.851	8.930
Step 2	Look up the 10% SNIP threshold for the Publication Year. <u>1.590</u>	<u>1.573</u>
Step 3	Is this Journal's SNIP (step 1) at least as large as the 10% SNIP threshold (step 2)? No	Yes

Question:
How do I calculate Scholarly Output?

Answer: Count the number of publications in the entity.

Scholarly Output = 6

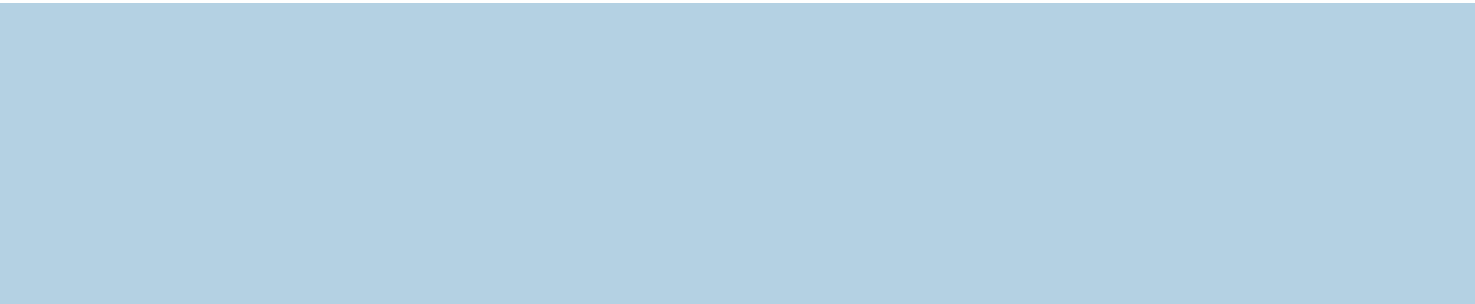
Question:
How do I calculate Publications in Top Journal Percentiles?

Answer: For "Absolute Number" display option, count the publications that are published in the top 10% of journals.

Publications in Top 10% Journals Percentiles = 2

– For "Percentage" display option, divide "Absolute Number" by the Scholarly Output of the entity.

Publications in Top 10% Journal Percentile = $(2/6) \times 100 = 33.3\%$



Entity with 6 Publications			
Publication 3	Publication 4	Publication 5	Publication 6
2007	2007	2006	2006
<i>Vision Research</i>	<i>Nature</i>	<i>Vision Research</i>	<i>Biophysics</i>
1.130	7.100	1.292	0.036
1.566	1.566	1.573	1.573
No	Yes	No	No

SNIP Percentile Thresholds			
Percentile Level	2005	2006	2007
1%	4.116	4.024	4.036
5%	2.089	2.047	2.034
10%	1.590	1.573	1.566
25%	1.072	1.061	1.045

4.16 Metric: *h*-indices

h-indices in SciVal indicate a balance between the productivity (Scholarly Output) and citation impact (Citation Count) of an entity's publications.

h-indices in SciVal offer 3 variants: the *h*-index, the *g*-index, and the *m*-index. The *g*- and *m*-indices inherit the positive qualities of the *h*-index, but address aspects that are sometimes considered shortcomings; therefore these metrics are grouped together into a set collectively called *h*-indices:

- *h*-index is now recognized as an industry standard that gives information about the performance of Researchers and Research Areas that is very useful in some situations. *h*-index of an entity is 9 if the top 9 most-cited publications have each received at least 9 citations; it is 13 if an entity's top 13 most-cited publications have each received at least 13 citations; and so on
- *g*-index is a variant of the *h*-index that emphasizes the most highly cited papers in a data set. The *h*-index does not give extra weighting to the most-cited publications of a data set that are likely the ones that are responsible for an entity's prestige; *g*-index can be used if this feature of the *h*-index is seen as a weakness. The *g*-index is always the same as or higher than the *h*-index
- *m*-index is another variant of the *h*-index that displays *h*-index per year since first publication. The *h*-index tends to increase with career length, and *m*-index can be used in situations where this is a shortcoming, such as comparing researchers within a field but with very different career lengths. The *m*-index inherently assumes unbroken research activity since the first publication

h-indices are available in SciVal for all Researcher-based entities, and for Research Areas.

h-indices are:

- Productivity metrics
- Citation Impact metrics
- Snowball Metrics (*h*-index only)
- “Power metrics”

A Researcher-based entity, or a Research Area, has a single value for the *h*-, *g*- or *m*-index that is based on all publications in the data set; these indices are not calculated per year. This is represented in the Benchmarking module as the same value for all years so that this metric can be made available in this module; in the Chart view, for instance, a horizontal line will be displayed and the years do not have a meaning.

These metrics are useful to:

- Benchmark activity in a way that relies on the **balance between 2 fundamental aspects of performance, namely productivity and citation impact:**
 - The total number of publications, or Scholarly Output, of an entity sets a limit for the value of the *h*-index. If a researcher has 1 publication that has been cited 100 times, their *h*-index cannot exceed 1
 - The total number of citations received, or Citation Count, sets the other limit for the value of the *h*-index. If a researcher has 100 publications which have each received 0 or 1 citations, their *h*-index also cannot exceed 1
- Be used as a **related group of metrics**, each with their own strengths
- **Provide transparency** on the underlying data to build trust in the metrics in SciVal

These metrics should be used with care when:

- **Comparing entities of significantly different sizes:**
 - The values of these metrics are limited by the Scholarly Output of an entity, and tend to increase with the size of the data set
 - This can be accounted for within a discipline, when the difference in size is due to different career lengths, by using the *m*-index; in this situation, variations revealed by the *m*-index are due to differences in annual productivity and citations received, which are likely the performance aspects of interest
- **Benchmarking entities within different disciplines**, even if these entities have similar sizes:
 - The values of *h*-indices are limited by the Citation Count of an entity, and tend to be highest in subject fields such as biochemistry, genetics and molecular biology; this reflects distinct publication and citation behavior between subject fields and does not necessarily indicate a difference in performance
 - It is not advisable to compare the *h*-indices of entities that fall entirely into distinct disciplines, such as a Researcher in genetics with a Researcher in human-computer interaction
 - When comparing entities made up of a mixture of disciplines, such as cross-disciplinary research teams, it is advised to apply the Research Area filter to focus on one field that is common between all the entities
- **An indication of the magnitude of the productivity and citation impact of an entity is important.**
It is advised to use Scholarly Output and Citation Count when it is important to communicate scale
- **Entities are small and there may be gaps in their output within the Scopus coverage:**
 - A single missing publication from a total of 3 or 4 will have a significant negative impact on apparent performance, whereas the effect of 1 missing publication from a total of 100 may be acceptable
 - The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher, or to limit the use of these metrics to comparing larger data sets where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison
- There is a **concern that excessive self-citations may be artificially inflating the values.** Users can judge whether the level of self-citations is higher than normal by deselecting the “Include self-citations” option
- **Investigating the research output of early-career researchers** who have, as yet, few publications, and have received a low number of citations due to the short time since publication. These metrics will not give reliable information in this situation, and it would be preferable to use metrics such as Collaboration or Publications in Top Journal Percentiles

Useful partner metrics are:

- **Scholarly Output and Citation Count** which provide information about the magnitude of productivity and citation impact
- **The set of all other “Power metrics”** whose value tends to increase as the entity becomes bigger: Scholarly Output, Journal Count, Journal Category Count, Citation Count, Cited Publications (“Absolute number”), Number of Citing Countries, Collaboration (“Absolute number”), Academic-Corporate Collaboration (“Absolute number”), Outputs in Top Percentiles (“Absolute number”), and Publications in Top Journal Percentiles (“Absolute number”)
- **The set of “time-independent metrics”** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Journal Count, Journal Category Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Example 9: *h*-indices

Scenario: The user would like to calculate the *h*-indices of an entity that consists of 10 publications. They have not selected any viewing or calculation options.

Entity with 10 Publications			
Publication Year	2004	2003	1998
Total citations received by this publication	2	5	4
Order and rank the publications			
Step 1	Sort the publications by their citation counts, from largest to smallest.	12	6
Step 2	Assign a rank to each publication, starting from 1.	1	3
Question:	How do I calculate <i>h</i> -index?		
Answer:	- Is the citation count for the publication (step 1) as least as large as its rank (step 2)?	12 ≥ 1 Yes	6 ≥ 3 Yes
	- Identify the highest value for which citation count is at least as large as the publication's rank.	<i>h</i> -index ≥ 1	<i>h</i> -index ≥ 3
Question:	How do I calculate <i>g</i> -index?		
Answer:	- Sum the citation counts of the publications ranked up to and including the current position.	12 12	12+12 24
	- Calculate the square of the rank (from step 2).	1x1 1	3x3 9
	- Is the sum of the citation counts at least as large as the square of the rank?	12 ≥ 1 Yes	24 ≥ 9 Yes
	- Identify the highest value for which the sum of the citation counts is at least as large as the square of the rank	<i>g</i> -index ≥ 1	<i>g</i> -index ≥ 3
Question:	How do I calculate <i>m</i> -index?		
Answer:	- Earliest Publication Year in set	1997	
	- Current Year	2014	
	- Calculate number of years since earliest publication.	2014 - 1997 + 1 18	
	- Divide <i>h</i> -index by number of years	5 / 18 <i>m</i>-index = 0.3	

Entity with 10 Publications						
1997	1997	1998	2008	1997	2001	2002
12	4	12	4	6	2	5
5	5	4	4	4	2	2
4	5	6	7	8	9	10
5 ≥ 4 Yes	5 ≥ 5 Yes	4 ≥ 6 No	4 ≥ 7 No	4 ≥ 8 No	2 ≥ 9 No	2 ≥ 10 No
<i>h</i> -index ≥ 4	<i>h</i>-index = 5	<i>h</i> -index < 6	<i>h</i> -index < 7	<i>h</i> -index < 8	<i>h</i> -index < 9	<i>h</i> -index < 10
12+12+6+5 35	12+12+6+5+5 40	12+12+6+5+5+4 44	12+12+6+5+5+4+4 48	12+12+6+5+5+4+4+4 52	12+12+6+5+5+4+4+4+2 54	12+12+6+5+5+4+4+4+2+2 56
4x4 16	5x5 25	6x6 36	7x7 49	8x8 64	9x9 81	10x10 100
35 ≥ 16 Yes	40 ≥ 25 Yes	44 ≥ 36 Yes	48 ≥ 49 No	52 ≥ 64 No	54 ≥ 81 No	56 ≥ 100 No
<i>g</i> -index ≥ 4	<i>g</i> -index ≥ 5	<i>g</i>-index = 6	<i>g</i> -index < 7	<i>g</i> -index < 8	<i>g</i> -index < 9	<i>g</i> -index < 10



For more information about SciVal,
please visit: elsevier.com/research-intelligence/scival

02.14 MKT