



Universiteit
Leiden
The Netherlands

Bibliographic analyses of handbooks: A proposal. Handbooks as a source of data for mapping scientific disciplines

Kroonenberg, P.M.; Costas, R.; Franssen, T.; Yegros, Yegros A.

Citation

Kroonenberg, P. M. (2018). Bibliographic analyses of handbooks: A proposal. Handbooks as a source of data for mapping scientific disciplines. *Cwts Conference Proceedings*, 1609-1618. Retrieved from <https://hdl.handle.net/1887/65225>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/65225>

Note: To cite this publication please use the final published version (if applicable).



STI 2018 Leiden

*23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"*

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at <https://hdl.handle.net/1887/64521>

ISBN: 978-90-9031204-0

© of the text: the authors

© 2018 Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands



This ARTICLE is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Bibliographic analyses of handbooks: A proposal.

Handbooks as a source of data for mapping scientific disciplines¹

Pieter M. Kroonenberg*

[*kroonenb@fsw.leidenuniv.nl](mailto:kroonenb@fsw.leidenuniv.nl)

Institute of Education and Child Studies, Leiden University, Wassenaarseweg 52, 2333 AK Leiden (The Netherlands); The Three-Mode Company, Wasstraat 11, 2313 JG Leiden (The Netherlands).**

Introduction

The aim of this paper is to sketch how one may use the references in handbooks to map scientific disciplines both in terms of their researchers and their research issues. This approach was taken in the study of attachment (Kroonenberg, Stoltenborgh and Mesman, 2010) based on the references from the first two editions of the Handbook of Attachment (Cassidy & Shaver, 1999, 2008). This paper promotes this approach but given the audience at STI conferences, the presentation of the methodology is based on the references from the *Handbook of Quantitative Studies of Science and Technology*, edited by Van Raan (1988).

However, the interpretation of the analysis outcomes for the scientometric handbook will be rather superficial and should be expanded. The primary purpose of this paper is to show that using bibliographical data from handbooks is a useful exercise for any scientific field.

Types of Handbooks

There are two kinds of “Handbooks”. The first kind is a reference book describing the state-of-the-art of a field of endeavour, often combined with a history of the subject and outlooks on its future. The second kind is a manual for some apparatus or for a style of behaviour containing instructions on how to behave or act to achieve a particular goal, say writing a thesis.

In this paper we will exclusively deal with the first kind. In particular, we will concentrate on analysing the references contained in the *Handbook of Quantitative Studies of Science and Technology* (Van Raan, 1988). In particular we will investigate (1) co-authorships of papers referenced in the Handbook (Who publishes with whom) and (2) who is referenced in which chapter (Who published what in which subfield?). In the second case the assumption is that if the same authors are referenced in different chapters, the content of these chapters must be in some way related and together they describe a particular subfield.

Co-authorships of papers are assumed to indicate that there is a working relationship between the authors either within or between research groups. Authors who have published with many different authors, especially with other key authors, are considered ‘spiders’ in the web that is formed by all co-authorships.

¹ The invaluable comments of Prof. Van Raan and the reviewers are gratefully acknowledged. Given the counting has been done by one person only once, some minor inaccuracies in the numbers may be present.

Handbooks: Descriptions of a Substantive Field

In this section we will give a short résumé of the stated aims of three handbooks to support the claim that they can be considered as a more or less complete description of the content of a discipline. The claim is strengthened by statements in their introductions that the chapters were written by eminent scientists in the field.

Handbook of Attachment: Theory, Research, and Clinical Applications

“Widely regarded as the state-of-the-science reference on attachment, this handbook interweaves theory and cutting-edge research with clinical applications. Leading researchers examine the origins and development of attachment theory”

<https://www.guilford.com/books/Handbook-of-Attachment/Cassidy-Shaver/9781462536641>

APA Handbook of Clinical Psychology

“The 5-volume APA Handbook of Clinical Psychology reflects the state-of-the-art in clinical psychology — science, practice, research, and training. The Handbook provides a comprehensive overview of: the history of clinical psychology, specialties and settings, theoretical and research approaches, [...], and future directions of clinical psychology.”

<http://www.apa.org/pubs/books/4311523.aspx> (Norcross, VandenBos, & Freedheim., 2016).

Handbook of Quantitative Studies of Science and Technology

“The principal purpose of this handbook is to present this wide range of topics in sufficient depth to give readers a reasonably systematic understanding of the domain of contemporary quantitative studies of science and technology, a domain which incorporates theory, methods and techniques, and applications.[..]. (Van Raan, 1988).

Authors and references

Given the situation sketched in the previous section, the following questions arise (1) how key authors of the research field covered by a handbook can be identified, (2) how their mutual relationships can be unearthed, and (3) how the major subject domains in the field can be extracted. In this paper, the major statistical instruments for these are, respectively: (1) a count of the number of times an author is referenced in the handbook, (2) a multidimensional scaling analysis of co-author information, and (3) a correspondence analysis of which authors are referenced in which chapters.

Many technical and graphical representations have been developed over the years which can enhance the presented analyses. The techniques used to analyse the handbook data are already well established in scientometrics. It is their conceptual use in analysing references from handbooks which is our prime focus. More subtle, especially graphical, procedures to enhance the analyses can, for instance, be found in the work of Van Eck and Waltman (2010, 2014).

Key authors (Who are key authors?)

Key authors are those authors who are most frequently referenced in a particular handbook. However, special care must be taken with authors who are the only ones referenced in a particular chapter but who have not been referenced in any other chapter.

Key authors are not necessarily the most referenced persons in the field as such, but they certainly have been recognised as having written works that authors felt they had to cite. In practice it was decided to designate as key authors those writers who appeared most often in

the author index of the handbook. There is a natural bias towards chapter authors due to self-citation, but given they are selected for their eminence by the editors, the number of entries in the author index seems an acceptable definition.

There are, of course, a few drawbacks and inherent weaknesses in this approach. To begin with, in the case of manual rather than digital counting, authors who were referenced several times on the same page, only get one entry in the author index. Moreover, authors get mentioned only once for their occurrence on a page in a chapter's reference list, irrespective of their number of references on that page. This is particularly unfortunate for first authors as they are clustered in a reference list. Practically this problem can only be eradicated if a searchable electronic copy of a handbook is used.

Author web and spiders (Who published with whom?)

The chapter references were scanned to see how often key authors had published papers with other key authors. The co-authorships were entered into a square matrix of key authors by key authors (for the principle, see Figure 1, left-hand panel; for an example see Table 1). This symmetric matrix is suitable for analysis by multidimensional scaling (see below).

In a multidimensional scaling plot a line can be drawn between each pair of co-authors and the collection of lines is referred to as the author web. Those authors are characterised as spiders, when they are co-author with many other authors. It is difficult to formulate a general rule for the number of connections that a key author needs to have to become a spider. Much will depend on the data themselves and the number of existing co-authorships.

Field structure (Who published what in which subfield?)

Each time a key author is mentioned in a chapter's reference list an appropriate tick is given in the key author by chapter matrix (see Figure 1, right-hand panel). Such a matrix can be transformed into a matrix with key author profiles, i.e. each row (author) is divided by its row total and into a matrix with chapter profiles, i.e. each column (chapter) is divided by its column total. If a row/column profile is similar to another row/column profile they are depicted close together in the biplot produced by a correspondence analysis (see below).

Data

The data structures of the two questions raised here (*who publishes with whom?* and *who publishes what in which subfield?*) are quite different and are outlined in Figure 1.

Multidimensional scaling

In multidimensional scaling co-author information is first transformed from similarities into distances, which are then analysed. The major outcome is an (often two-dimensional) plot in which co-authors are placed close to each other; small distances mean often published together, large distances mean never or rarely published together. Details about multidimensional scaling can, for instance, be found in the book by Borg, Groenen and Mair (2018).

Correspondence analysis

The contingency table with the information which authors are referred to in which chapters will be analysed by correspondence analysis. The technique handles both chapter profiles and author profiles (see right-hand panel Table 1). Both types of profiles can be displayed separately as well as simultaneously in a two-dimensional graph. In the simultaneous graph, a biplot, the relationships between authors and chapters can be read from the projections of the

author points on the chapter arrows, referred to as biplot axes, or vice versa. The chapter positions in the plot with respect to each other provide a condensed overview of the reference-based organisation of the field. Note that the dimension axes in correspondence analysis plots are mathematical constructs and need not necessarily themselves be interpreted. Details about correspondence analysis and its biplots can be found in Greenacre (1985, 2016) and many other books on the subject, such as Beh & Lombardo (2014).

Figure 1. Data formats for the two basic questions of this paper.

Left-hand panel: Who published with whom?

Right-hand panel: Who published what in which subfield?.

Research question	<i>Who published with whom?</i>	<i>Who published what?</i>																																										
Data format	<p>Similarity matrix of authors by authors</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="3">Authors</th> </tr> <tr> <th>A</th> <th>B</th> <th>C</th> </tr> </thead> <tbody> <tr> <th rowspan="3">Authors</th> <th>A</th> <td>x</td> <td>f_{AB}</td> <td></td> </tr> <tr> <th>B</th> <td>f_{BA}</td> <td>x</td> <td></td> </tr> <tr> <th>C</th> <td></td> <td></td> <td>x</td> </tr> </tbody> </table> <p>$f_{BA} = f_{AB}$ = number of times papers written by (co-)authors A and B have been cited in an entire <i>Handbook</i>.</p>			Authors			A	B	C	Authors	A	x	f_{AB}		B	f_{BA}	x		C			x	<p>Cross-tabulation of authors by chapters</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="3">Chapters</th> </tr> <tr> <th>1</th> <th>2</th> <th>3</th> </tr> </thead> <tbody> <tr> <th rowspan="3">Authors</th> <th>A</th> <td></td> <td></td> <td></td> </tr> <tr> <th>B</th> <td></td> <td>f_{B2}</td> <td></td> </tr> <tr> <th>C</th> <td></td> <td></td> <td></td> </tr> </tbody> </table> <p>f_{B2} = number of papers of which author B was a co-author cited in Chapter 2.</p> <p>(f_{A2}, f_{B2}, f_{C2}) is profile of Chapter 2;</p> <p>(f_{B1}, f_{B2}, f_{B3}) is profile of Author B</p>			Chapters			1	2	3	Authors	A				B		f_{B2}		C			
				Authors																																								
		A	B	C																																								
Authors	A	x	f_{AB}																																									
	B	f_{BA}	x																																									
	C			x																																								
		Chapters																																										
		1	2	3																																								
Authors	A																																											
	B		f_{B2}																																									
	C																																											
Analysis technique	Multidimensional scaling	Correspondence analysis																																										

Handbook of Quantitative Studies (Van Raan, 1988).

Key authors (Who publishes with whom?)

From the author index a number of key authors were identified. There were 15 authors with 23 or more entries in this index; however L. Simon who had 24 entries was excluded as there were only references in one single chapter written by Franklin as co-author. For the 14 key authors it was counted how many times they co-authored papers with each other (see Table 1).

Author web

The results were (for me) rather unexpected as the number of co-authored papers by key authors was rather limited. Moreover, some further inspection of the origins of co-authors brought to light that only those key authors wrote papers together when they belonged to the same research group. In particular, Van Raan, Moed, and Frankfort all worked at CWTS, Leiden, The Netherlands. The brothers Stephen and Jonathan Cole both worked at Columbia University, New York and were part of a group around Merton. Small and Griffith both

worked at Garfield's Institute of Scientific Information in Philadelphia, and both Narin and Carpenter worked at Computer Horizons, Inc.

Because the groups published only in isolation (the handbook was the first time they all appeared in a volume together) multidimensional scaling could not be meaningfully applied due to the lack of links between the groups. For a meaningful presentation of the authors in a multidimensional scaling graph co-authorships between the groups should be present.

Using multidimensional scaling, Figure 2 gives an impression of a co-author web in which many key authors are connected with each other through common publications. The example is taken from our attachment report (Kroonenberg, Stoltenborgh, & Mesman, 2010). The co-authors form a network indicated by the heavy lines; to emphasise this, the connections between key authors and other authors have been deleted.

A possible reason for the difference in connectedness could be that in 1988 the field of scientometrics was still relatively young, and no world-wide co-authoring network of the most important research groups had yet evolved, while this was the case for the field of attachment in 1999.

Table 1. Co-authorships between the ten most referenced authors. Basis for the "Who publishes with whom?"-question

	Author Index Entries	Cole JR	Cole S	Van Raan	Moed	Frankfort	Garfield	Small	Sweeney	Griffith	Narin	Carpenter	Rip	Franklin	Da Solla Price
Cole, JR	27		13												
Cole, S	27	13													
Van Raan, AFJ	42				14	11									
Moed, HF	27			14		11									
Frankfort, JJ	23			11	11										
Garfield, E	41							3							
Small, HG	60						3		13	8					
Sweeney, E	23							13							
Griffith, BC	31							8							
Narin, F	50											15			
Carpenter, MLP	24										15				
Rip, A	26														
Franklin JJ	27														
Da Solla Price, DJ	23														

Field structure (Who published what in which subfield?)

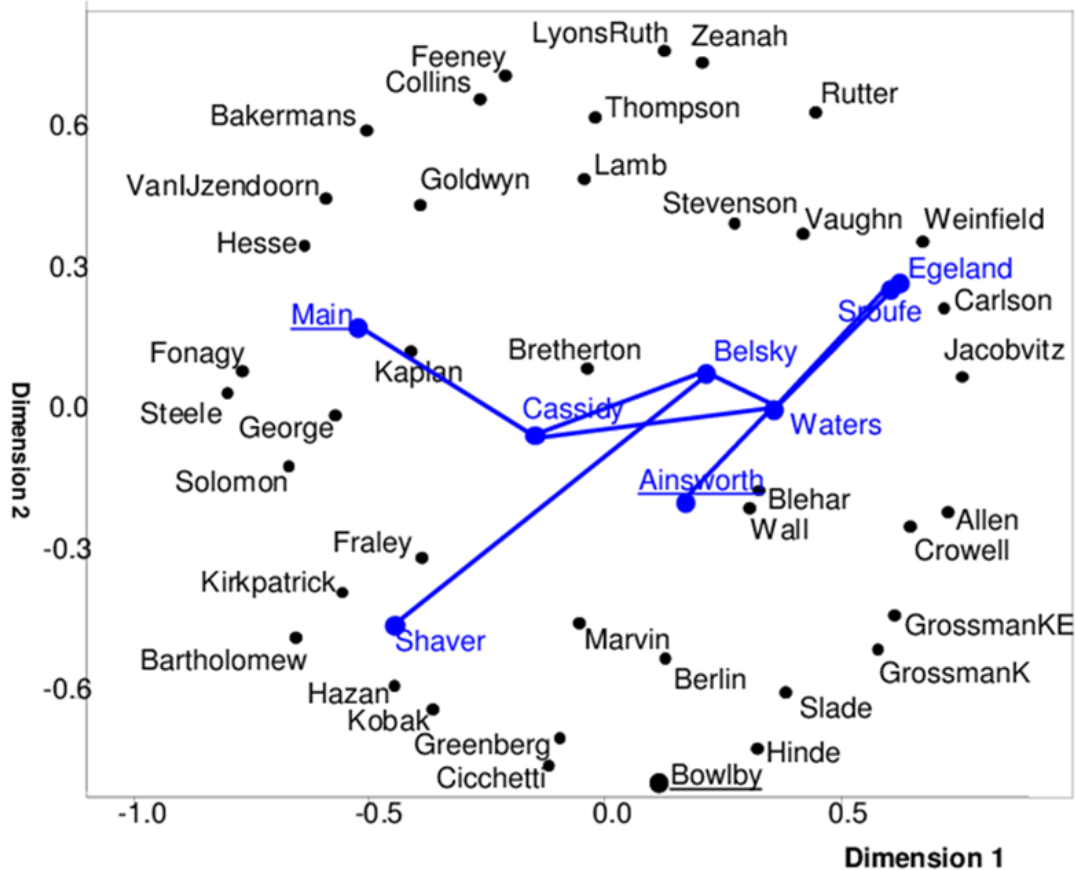
To evaluate who was working in which subfields and which chapters are pertaining to the same subfield, a correspondence analysis was applied to the Author-by-Chapter matrix containing the number of references in each chapter to the 42 most referenced authors (12 entries or more in the author index).

It turned out that the reference lists of Chapters 18 and 19, both dealing with technology, contained only papers of one single author of our chosen 42; i.e. Saviotti, who not only wrote one of the chapters (18) but was not referenced in any of the other chapters. This means that neither the chapters nor Saviotti could be placed in relation to the other chapters or authors and they were therefore not included in the correspondence analysis.

The basic interpretation of biplots such as Figures 3 (Dimension 1 and 4) and 4 (Dimensions 2 and 3) is portrayed in Figure 5. Each author can be projected onto each of the chapters; this is shown for Narin and Van Raan with respect to the chapters dealing with patent analysis (Chapters 15 and 16);. The arrows plus their extensions to the opposite side are mostly called biplot axes. A projection on the positive side like that of Narin indicates that Narin is more often than average referenced in the lists of the patent chapters. A projection on the negative side like that of Van Raan indicates that he is cited less than average in those chapters. Note

that it is not the distance between a point and a biplot axis which is relevant but only where on the axis the point's projection lies.

Figure 2. Attachment study: Who published with whom? Co-authorships of key authors



Field structure (Who published what in which subfield?)

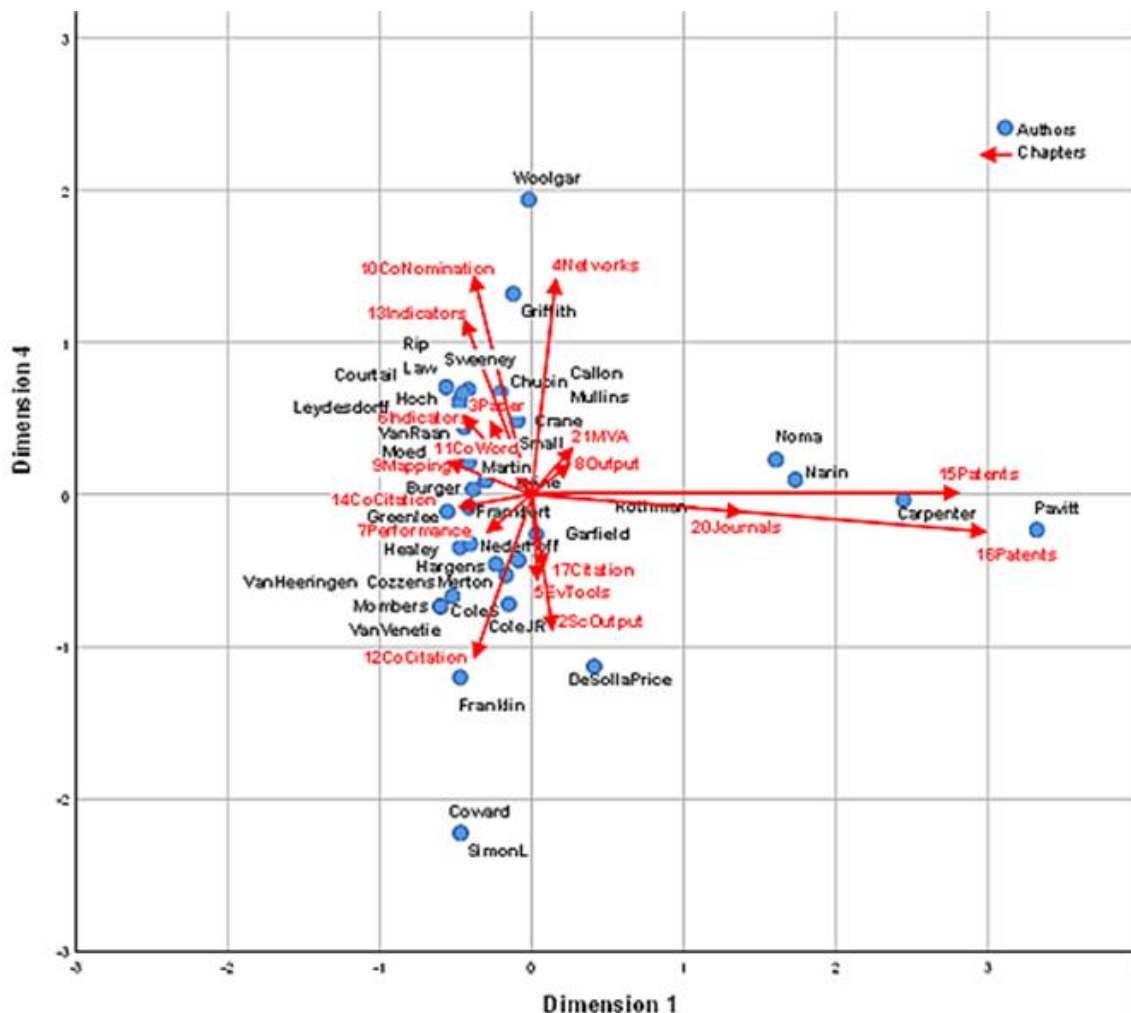
The first dimension in Figure 3 (Dimensions 1 and 4) shows that Chapters 15 and 16 dealing with patents refer largely to the same authors, especially Carpenter, Pavitt, Narin and Noma. It does not come as a surprise that Narin (Chapter 15) and Pavitt (Chapter 16) were co authors of these chapters.

The fourth dimension is characterised by a contrast of the Chapters 4 (Networks), 10 (Co-nomination) and 13 (Indicators) with especially Chapter 2 (Scientific Output) and 12 (Co-citation). In the latter chapters reference is especially made to Franklin, Simon, and Coward. Simon and Coward are exclusively referenced in Chapter 12 on co-citation, and Franklin almost exclusively so, reflecting their primary interest in co-citation analysis. One of the most important founding fathers of scientometrics, De Solla Price, is not often referenced, but his references are spread over in a number of different chapters. Woolgar is only referenced 5 times in reference lists, three times of which were in Chapter 4 on networks.

Even though the correspondence analysis shows who is mentioned in the reference list of which chapter, it is important to realise that occurrences in single chapters but no or few

references in other chapters can make authors stand out in the correspondence analysis plot. In Figure 3 Simon, Coward and Woolgar are examples of this.

Figure 3. Correspondence analysis: Who published what in which subfield?



Notes: Authors are indicated by their full name and chapters by the chapter number followed by a subject indication. Of a four-dimensional solution which accounts for 66% of the information in the data, the combination of the first and fourth dimension were chosen for this figure, because it provided a relatively clear picture of particular relationships between authors and chapters.

The same technique of interpretation as used for Figure 3 (Portraying dimensions 1 and 4) can be employed for Figure 4 (Portraying dimensions 2 and 3). Note that the two figures refer to different dimensions and therefore describe different patterns. The details and explanations of the patterns displayed will, however, will not be presented here, but left for the scientometrics experts.

Figure 4. Correspondence analysis: Who published what in which subfield?

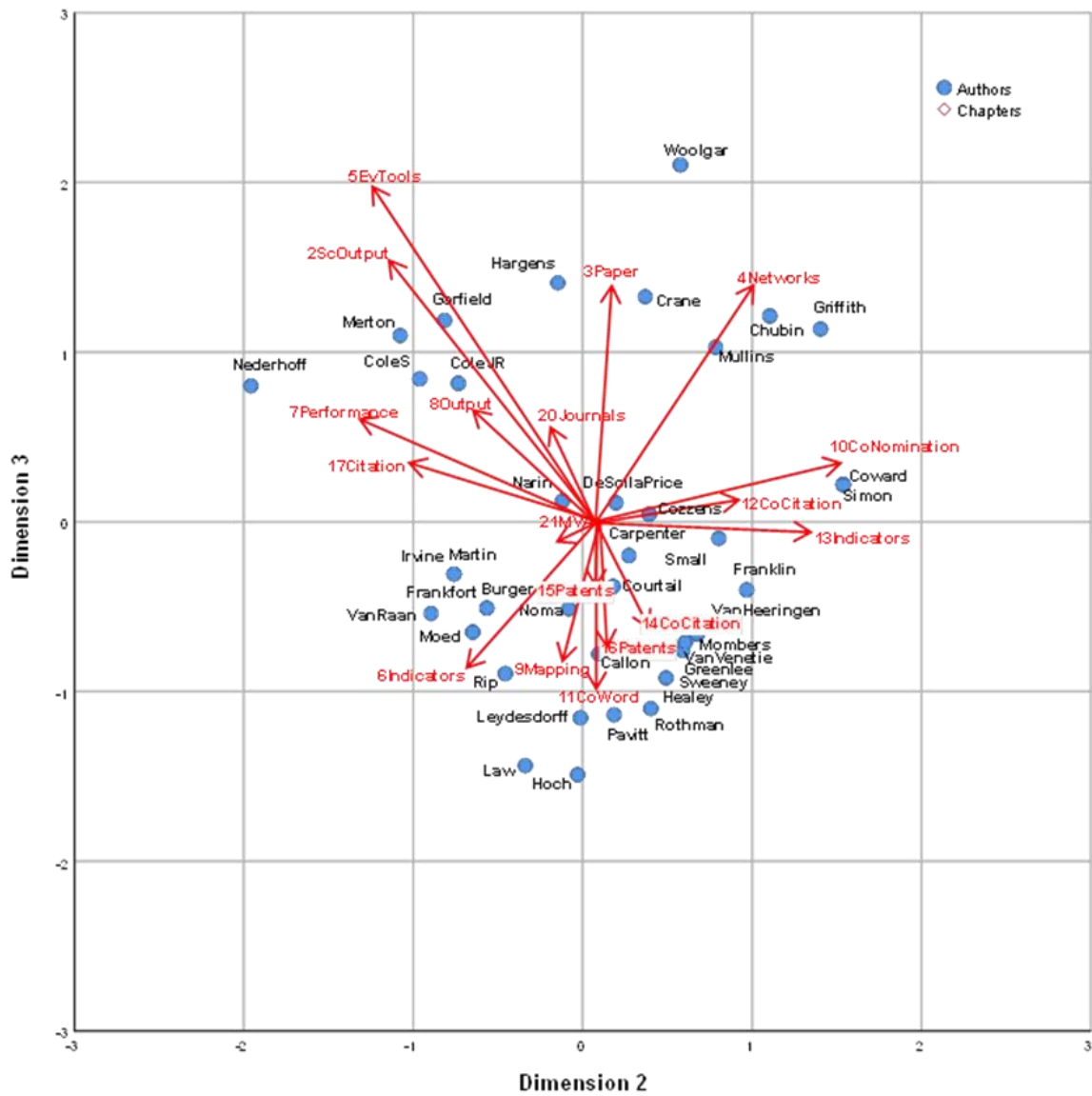
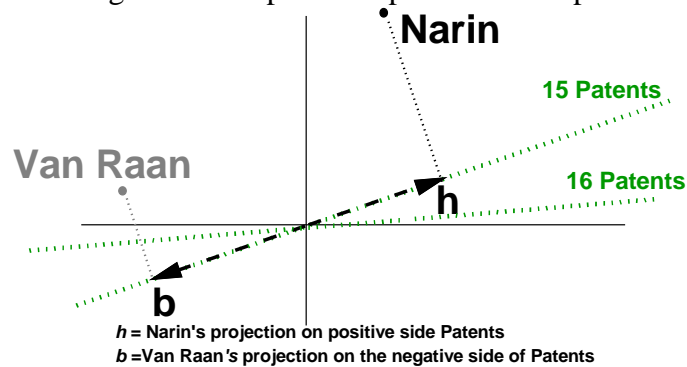


Figure 5: Interpretation protocol for biplots.



Note: The example axes represent Chapters 15 and 16, both dealing with patent analyse

Conclusion

In this paper a case is made for using the bibliometric information contained in handbooks for examining the structure of a discipline, both in terms of its subfields and who works in the area with whom. This proposal was applied in detail in the report *The World of Attachment* (Kroonenberg, Stoltenborgh, Mesman, 2010).

To call the attention to this proposal, a rudimentary analysis was made of the scientometrics *Handbook of Quantitative Studies of Science and Technology* edited by Van Raan in 1988. The emphasis was on procedural matters rather than on content, and it is clear that further interpretational details can be worked out. Moreover, several sophisticated graphical aids already described in the literature and embedded in computer programs can be usefully employed for the handbook data (see the work of Van Eck and Waltman, 2010, 2014). However, this can only be done in cooperation with experts from the field who can assess the outcomes based on their substantive knowledge.

What is lacking in this paper as well is the embedding of the proposal into the standard bibliometric literature, but this is an investigation for the future. Such a study shall have to show which aspects are new and which are already available in the literature.

References

- Beh, E. J. & Lombardo, R. (2014). *Correspondence analysis. Theory, Practice and New Strategies*. Chichester: Wiley.
- Borg, I., Groenen, P.J.F., & Mair, P. (2018) *Applied Multidimensional Scaling and Unfolding*. New York: Springer.
- Cassidy, J. & Shaver, P.R. (Eds.) (1999). *Handbook of attachment. Theory, Research, and Clinical Applications*. New York: The Guilford Press.
- Cassidy, J. & Shaver, P.R. (Eds.) (2008). *Handbook of Attachment. Theory, Research, and Clinical Applications* (2nd ed.). New York: The Guilford Press
- Greenacre, M.J. (1985). *Theory and Application of Correspondence Analysis*. London: Academic Press.
- Greenacre, M.J. (2016). *Correspondence Analysis in Practice* (3rd edition). Boca Raton, FL: Chapman & Hall/CRC
- Kroonenberg, P.M., Stoltenborgh, M., & Mesman, J. (2010). *The World of Attachment seen through the eyes of the Handbooks of Attachment*. (Available from: <https://openaccess.leidenuniv.nl/bitstream/handle/1887/16294/HandBookReport2010.pdf?sequence=1> also via ResearchGate.net and Academia.edu)

Norcross, J.C., VandenBos, G.R., & Freedheim, D.K. (Eds.) (2016). *APA Handbooks in Psychology. APA Handbook of Clinical Psychology*. Washington, DC: American Psychological Association.

Van Eck N.J., L. Waltman (2010). Software survey: VOSviewer, A computer program for bibliometric mapping. *Scientometrics*, 84, 523–538. <http://www.vosviewer.com/home>.

Van Eck N.J., L. Waltman (2014). CitNetExplorer: A new software tool for analyzing and visualizing citation networks. *Journal of Informetrics*, 8, 802-823. <http://www.citnetexplorer.nl/home>.

Van Raan, A.F.J. (Ed.) (1988). *Handbook of Quantitative Studies of Science and Technology*. Amsterdam: North Holland.