

Using bibliometric maps of science in a science policy context

Ed Noyons

ABSTRACT

Within the context of science policy new softwares has been created for mapping, and for this reason it is necessary to reflect on standardized methods and their application in bibliometric analysis. This paper presents the method named Visualization of Similarities (VOS) and VOSviewer application, developed by Erasmus University and Leiden University, and describes the process of creating maps and how they can be used in the context of management and science policy. The method and software are used in around 40,000 published documents by Brazil and the Netherlands in the area of Brain and Cognition research in the period 2001 to 2011. The maps created show the area growth, the relationships between the subareas of knowledge and between authors' institutions affiliation.

KEYWORDS: Mapping Science. Networks. VOSviewer.

1 Introduction

Maps of science, particularly those using bibliographic data, have become more popular in the past decades. There is substantial increase of bibliometric maps of science thanks to more data becoming available, computer becoming more powerful and technology becoming more useable in this context.

Within the context of science policy, we also see an increase of bibliometric mapping. More software is available to create maps of science (fields) and to use them as an interface to explore. For that reason, this seems to be a good time for some kind of standardization of method and application. Because in this area method and application are so intermingled, it is hard to treat them separately. In this paper I will try to contribute to the standardization of the application of bibliometric maps in the science policy context. Although I will try to deal with this aspect only, I realize that a reference to technique and method is often needed. The method I use in this paper is developed at Erasmus University and Leiden University by Nees Jan van Eck and Ludo Waltman (ECK ; WALTMAN, 2009, 2011 ; WALTMAN et al., 2010). The Visualization of Similarities (VOS) as a method to structure data and the VOSviewer as an application have proven to be very powerful but will not be discussed as such in this paper. Only the output of the method and technique will be used in this study.

The main objective of this paper is to describe the process of creating a map in a generic procedure and to show how they can be used in a science policy and research management context. The choices I make in my approach are based in Noyons (1999) plus the experiences at CWTS in the past decades in this context. I will use a recent case study in the field of Brain & Cognition Research (BCR) in which around 40, 000 publications were collected in the period 2001-2011 combined with all publications with a brazilian or dutch address within this set.

Regarding the use of maps in a science policy context we aim to address questions such as: what are the main developments in a field; what is the position of a certain actor (country, institute, etc.) in that field; how do these relate.

2 Data collection and methodology

BCR is a multi-disciplinary field studying the brain as well as its 'application areas'. As such it is an interesting field to be studied by the approach we discuss in this paper because it is not represented by traditional classifications.

The basic idea behind this multi-disciplinary field can be found on the website of the Dutch *National research initiatief Hersenen and Cognitie* (NIHC), a BCR program of the Netherlands Research Council (NWO).

One of the biggest challenges in today's brains and cognition research is to integrate different disciplinary approaches to understand specific cognitive processes, such as consciousness, language or memory from the molecular level to the level of human behavior and its surroundings. (NATIONAAL..., 2012, our translation)

Within this challenge many disciplines and fields are involved as we will see later. The fundamentals of BCR are discussed in international scholarly publications in journals mostly covered by the Web of Science, which makes a bibliometric study using the WoS feasible.

There is a substantial output produced by researchers in BCR but it appeared difficult to be isolated using a top-down approach. The publications defining the field do not 'automatically' cluster together. Therefore, we chose a bottom up approach to delineate BCR.

The bottom up approach basically works as follows: we first collect core publication data from the field and expand the core by collecting publications closely related to the core. The relatedness may be defined in different ways but for this study it was defined by direct citing relations. Core publications are publications which belong undeniably to the field. Such publications are typically identified by journal names or keywords.

The core publications set for BCR was determined by seven journals, unanimously suggested by three BCR experts. These journals are:

- a) Journal of Cognitive Neuroscience;
- b) Journal of Experimental Psychology-General;
- c) Journal of Neuroscience;
- d) Neuroimage;
- e) Neuron;
- f) Psychological Review;
- g) Psychological Science.

In a first iteration, we collected those publications (not in the core) that had at least 25 connections with the core (citing the core or being cited by the core). The expanded core was expanded once more in a second iteration by those publications

with at least 35 connections with the expanded core. In the second iteration we apply a higher threshold to be added to the core because the core expanded in the first iteration allows more peripheral publications to enter.

The delineation process (2 iterations) yielded almost 46,000 publications in 2001-2010. We use period of 10 years to have more robust data and to enable trend analyses.

It should be noted that the approach is biased towards highly cited papers and as such the data collection BCR used in this study is a sample of papers with a high impact in the field. For the purpose of this study we do not consider this a drawback.

Table 1 -- Number of selected publications in BCR (world-wide and for the Netherlands and Brazil)

Year	No of pubs	NL pubs	BR pubs
2001	3,876	73	12
2002	4,133	82	18
2003	4,442	120	10
2004	4,453	112	12
2005	4,525	134	19
2006	4,921	155	19
2007	4,934	169	27
2008	4,875	174	26
2009	4,802	197	28
2010	4,772	223	29

The amount of publications worldwide in BCR increases until 2007 and stabilizes since then. The production both in Brazil and the Netherlands steadily grows during the entire period. Does this mean Brazil and the Netherlands don't follow the general trends in BCR? We will get back to this issue in the results.

For different purposes in this study we needed to structure the entire set of BCR publications into relevant clusters to represent sub-domains of BCR. First of all, we wanted to identify sub-domains to understand the field better. Secondly we wanted to investigate in more detail the development of the field and finally we wanted to use the structure to characterize the contribution of different actors (e.g., Brazil, the Netherlands).

Structuring a field as defined by publications means identifying clusters of publications that are related. The relatedness can be based on different information but in fields such as BCR, citing relations are very convenient. BCR is well covered within WoS and the average number of references in the field is high (approximately 75 per publication). In this study, we created a network of all publications (almost 46,000) and their mutual citing relation. E.g.,

- Pub A cites Pub B
- Pub A cites Pub C
- Pub B cites Pub D
- Etc.

This entire network is analyzed using a method described in Waltman, van Eck and Noyons (2010), which yielded a structure in which we identified 31 clusters. Each of these clusters was labeled with a name by a BCR expert. The table below lists these 31 clusters with the number of publications included.

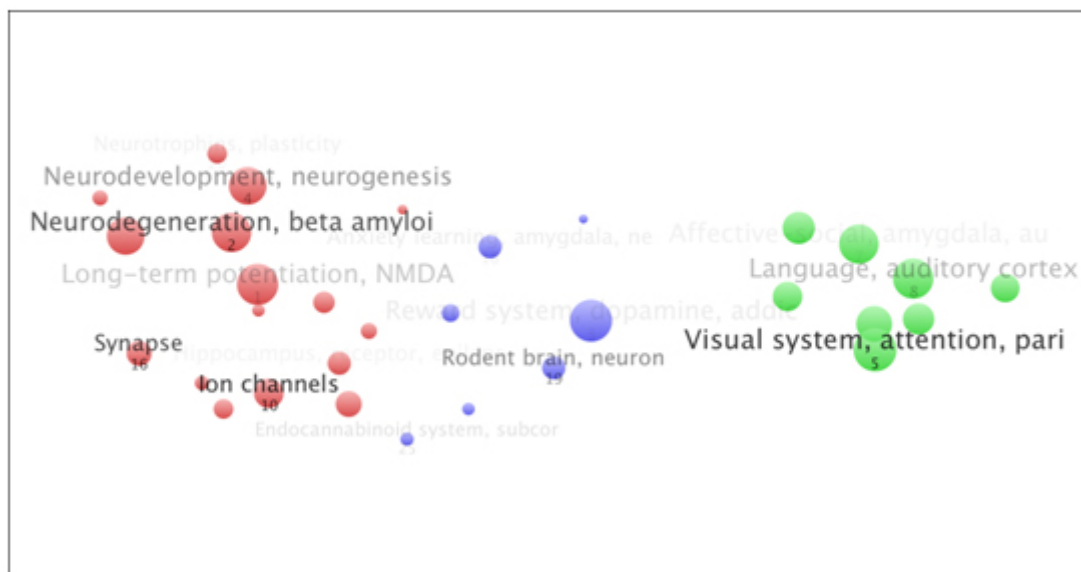
Table 2 -- Overview of sub-domains identified in BCR and their size (2001-2010)

Id	Label	No of pubs
1	Long-term potentiation, NMDA	3132
2	Neurodegeneration, beta amyloid, rodent model	2911
3	Reward system, dopamine, addiction	3240
4	Neurodevelopment, neurogenesis, stem cell	2564
5	Visual system, attention, parietal cortex	3395
6	Regeneration, axon, spinal cord	2615
7	Affective-social, amygdala, autism	2843
8	Language, auditory cortex	3031
9	MRI, connectivity	2397
10	Ion channels	1592
11	Memory, medial-temporal lobe	1650
12	Cognitive control, anterior cingulate, ADHD	1560
13	Motor system, TMS	1803
14	MRI, neurology, aging	1957
15	Pain, nerve injury	1289
16	Synapse	1187
17	Hippocampus, receptor, epilepsy	1012
18	Anxiety learning, amygdala, neurochemistry	1017
19	Rodent brain, neuron	1048
20	Neurotrophins, plasticity	753
21	Olfactory system, glutamate	845
22	Glia, glutamate	726
23	Sleep, neurochemistry	584
24	Synaptic plasticity, mental retardation	417
25	Endocannabinoid system, subcortex, neurotransmitters	338
26	Retina	406
27	Circadian clock, suprachiasmatic nucleus, intra-cellular signaling	302
28	Spinal cord, synapse	468
29	Acetylcholine, nicotine, synapse	316
30	Sex hormones, neurotransmitters	192
31	"Songbird rest"	133

The sub-domains as we identified them represent the structure field on a certain (locally optimal) resolution. This resolution may be too high for certain tasks but to illustrate the issues we want to raise in this paper it is suitable.

We enhance the structure (31 sub-domains) by visualizing their relation among each other in a map, a 2-dimensional representation. In this map the citing relations among the 31 sub-domains is used to position them in such a way that sub-domains that cite each other many times are in each other's vicinity while those who hardly cite each other are remote. By creating this map, we get more insight in the relations among sub-domains and we can observe the high resolution structure from a distance revealing the coarse structure.

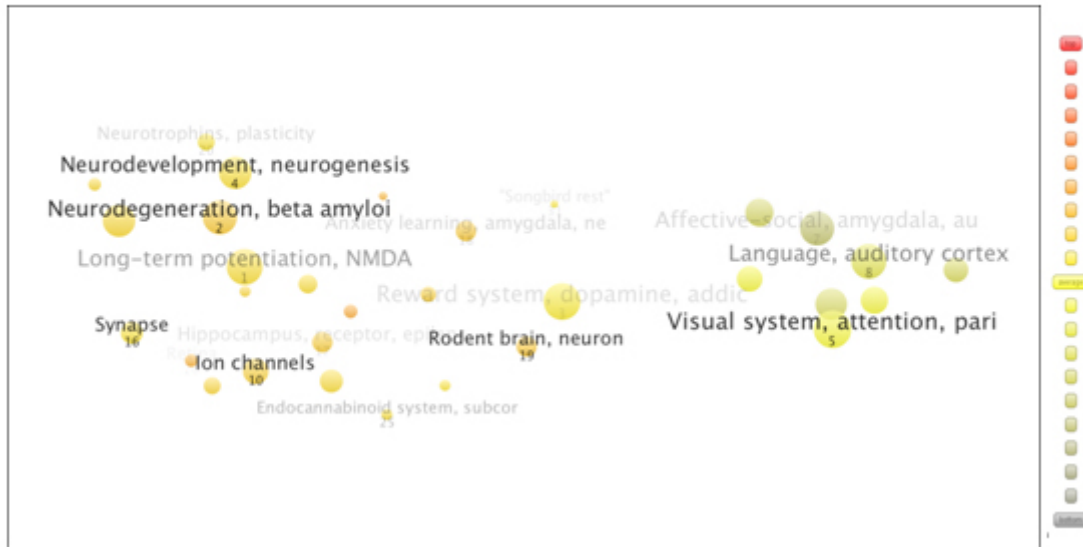
Figure 1 -- Bibliometric map of BCR 2001-2010



In the map the interaction among the 31 sub-domains (citation traffic) reveals an overall structure with three areas: the left hand area (Red) with almost half of all sub-domains, the right-hand area (Green) with almost one third of all sub-domains and a limited amount of sub-domains in the interface area (Blue).

The interpretation of the basic (an more detailed) structure is something that needs to be done in collaboration with the experts. There are, however, some possibilities to explore the map using additional bibliometric data. As all WoS publications are labeled with a subject category (through journal classification), we can show the distribution of the most important subject categories over the map. For instance the category 'Neurosciences' are distributed as depicted in Figure 2. In the entire BCR field there is a certain proportion of neurosciences papers. This is the expected proportion for each individual sub-domain. In the map we visualize the distribution neurosciences papers on the basis of the actual proportion in each sub-domains relative to the expected proportion.

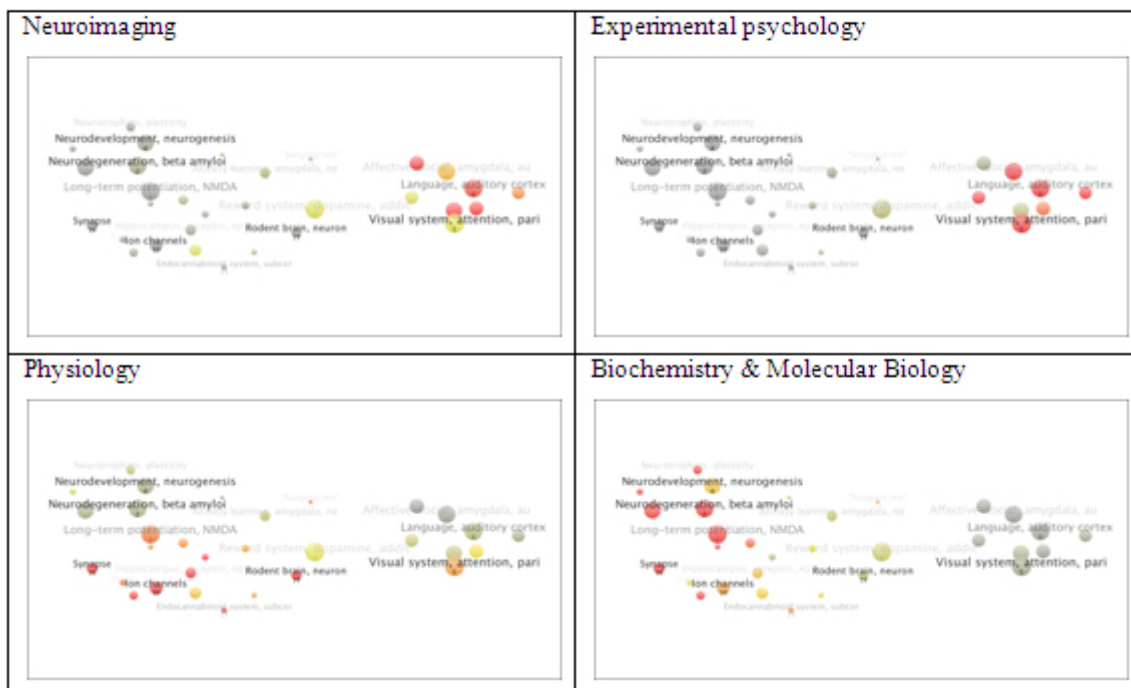
Figure 2 -- Distribution of Neuroscience papers over the BCR map



The color coding used in this map is shown in the legend, where yellow sub-domains are those in which the proportion neurosciences papers is as expected, while in the Red sub-domains the proportion neurosciences papers is above the expected. In this case it is clear that there is an even distribution of neurosciences papers over the entire map. All sub-domains in the map relate to neurosciences.

Subsequently, the categories *neuroimaging*, *experimental psychology*, *physiology* and *Biochemistry & Molecular biology*, respectively are plotted on the BCR map.

Figure 3 -- Distribution of publications in the 4 subject categories over BCR map



The overview of four distributions helps us to interpret the basic structure and provides a tool for the expert to do this. The right-hand area represents **neuroimaging** research as well as **psychology**. The experts call this area **cognitive neuroscience**. The left-hand area represents **physiologic research** as well as **biochemistry**, i.e., the **basic neuroscience** area. There is no particular category overrepresented in the interface area and as such difficult to label (also for the expert). Therefore we call this the **interface** area.

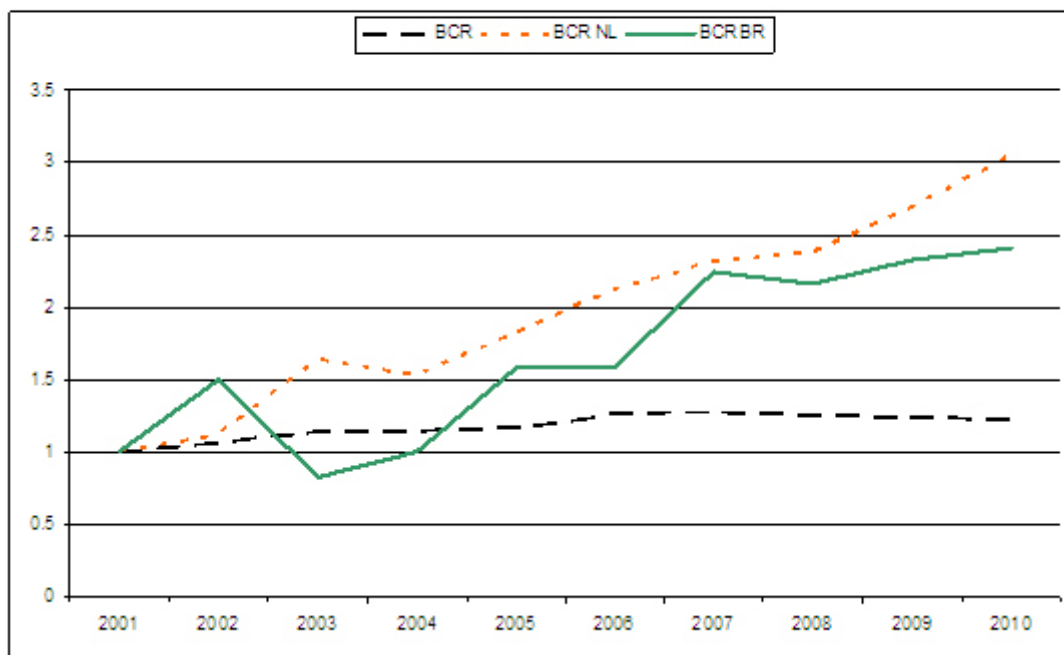
This method of visualization can also be used for actors, e.g., a country. We determine the proportion of a country in BCR as well as in each sub-domain and distribute the different proportions over the map, using a similar color-coding.

3 Results

We will discuss the potential of our approach using two similar case studies: the positioning and development of Brazil and the Netherlands in BCR.

In the previous section we discussed the data collection as well as the development of production world-wide in BCR in ten years. In Brazil and the Netherlands the development of production differed from the development world-wide. This is illustrated by the chart below.

Figure 4 – Yearly Growth of production indexed to 2001 in BCR (world-wide, Netherlands and Brazil)

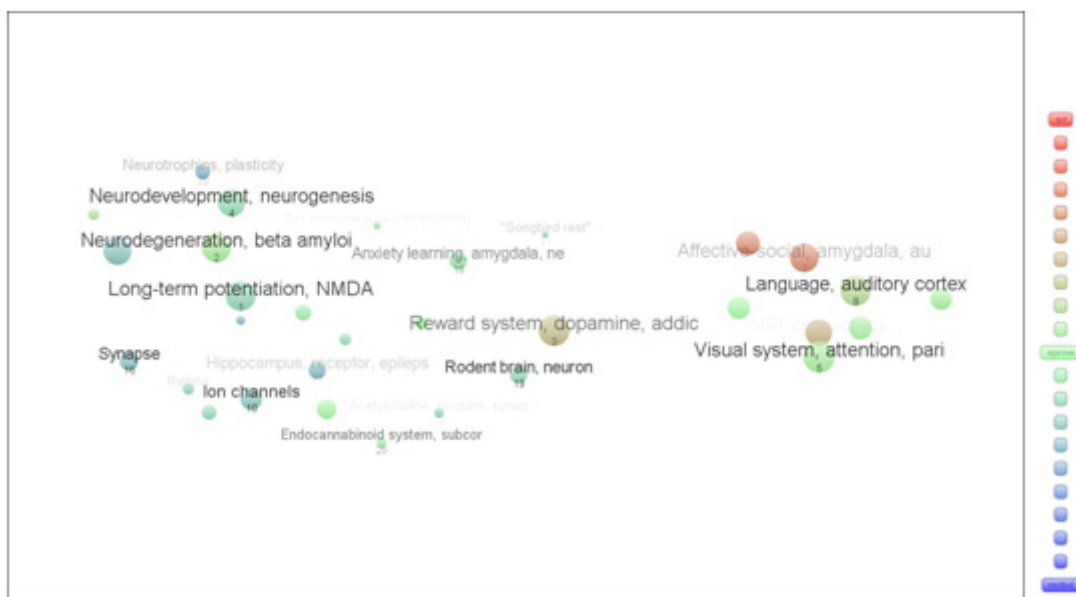


The picture shows for both Brazil and the Netherlands more production growth in ten years than world-wide in BCR. Does this mean that these two countries continue to publish more in a field in which the world has already lost its interest? By using

the map structure, we will show that is not the case.

First of all we use the map to monitor the developments in more detail. The entire field BCR shows a stabilized production over the studied period but the individual sub-domains all have their own development in terms of publication production. If we consider the BCR growth overall as the expected growth, we can measure the growth of each individual sub-domain compared to the expected growth. The degree to which the growth of an individual sub-domain exceeds the expected growth can be measured in a value, which is captured in a color (Blue-Green-Red). The result of this exercise is depicted in the map below.

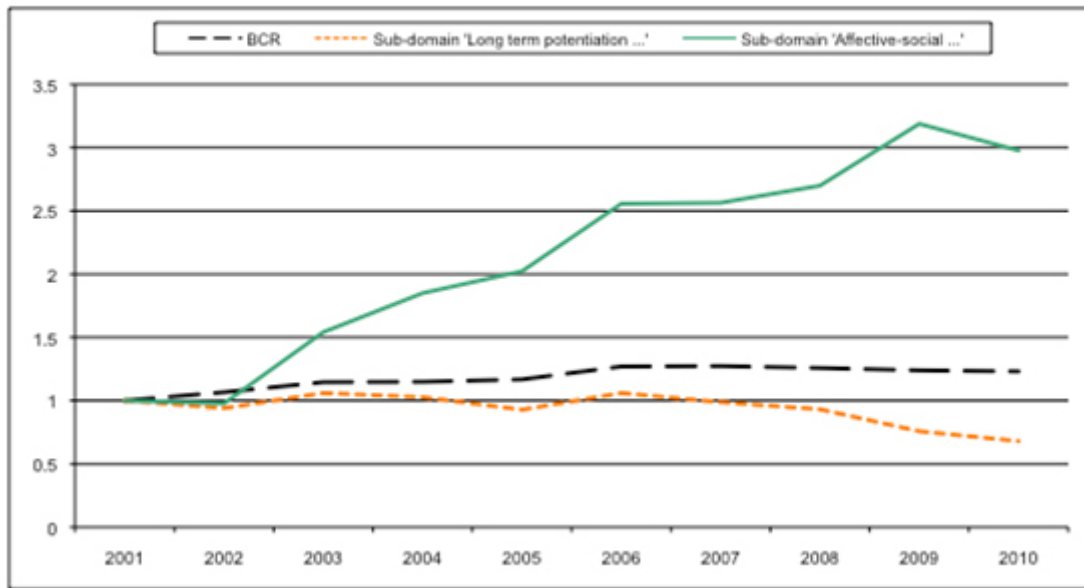
Figure 5 -- Relative growth of sub-domains in BCR



This map shows that the area of **cognitive neuroscience** shows a significant different development than BCR overall and the area of basic neuroscience. Almost all sub-domains in the cognitive neuroscience area show an increase exceeding world-wide BCR development. As an illustration we plotted the development of a sub-domain from the basic neuroscience area 'Long term potentiation, NMDA' and of a sub-domain from the cognitive neuroscience area 'Affective-social, amygdala, autism' in the chart below (Figure 6).

The trends reveal the remarkable difference of development between the basic neuroscience area and the cognitive neuroscience area. In the latter the production triples while in the former the production even shows a small decrease.

Figure 6 -- Yearly growth of two BCR sub-domains (indexed to 2001)

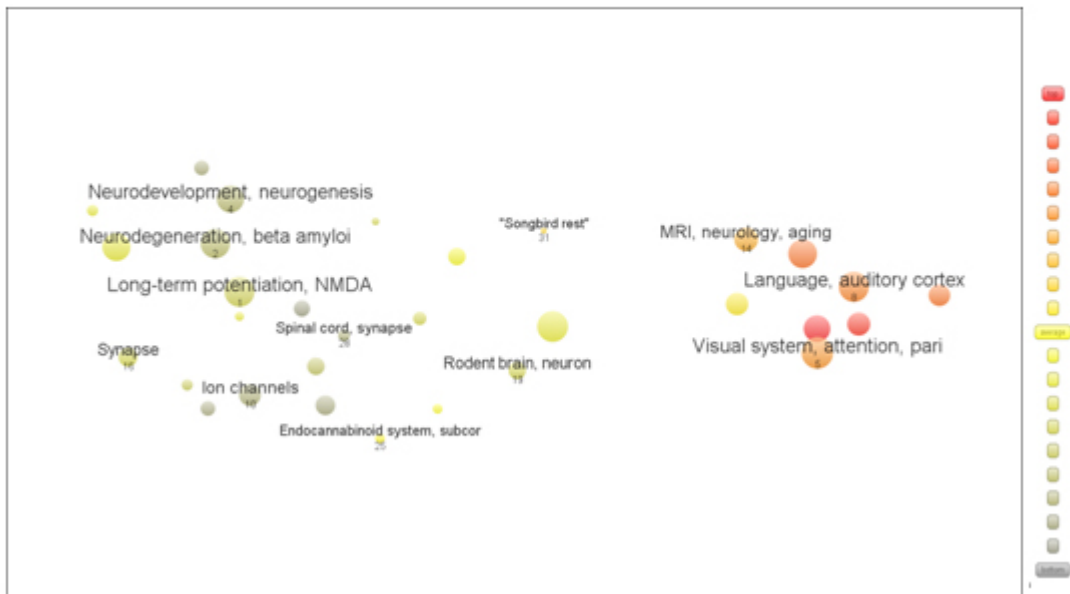


3.1 Case study #1 The Netherlands in BCR

As the a Dutch funding body in the Netherlands in the field of BCR wanted to know more about the developments in the field and the position of the Netherlands, we created an overview of Dutch production in BCR using the map.

We apply a similar color-coding as before to indicate the Dutch activity. We counted the total amount of BCR papers for the Netherlands and the number in each sub-domain. Thus we could determine the relative contribution of the Netherlands both in BCR and in each individual sub-domain. If we consider the Dutch proportion in BCR as the expected contribution, we can calculate the contribution in each sub-domain as a normalized contribution. The results are plotted in the BCR map.

Figure 7 – Relative distribution of BCR publications from the Netherlands (2001-2010)

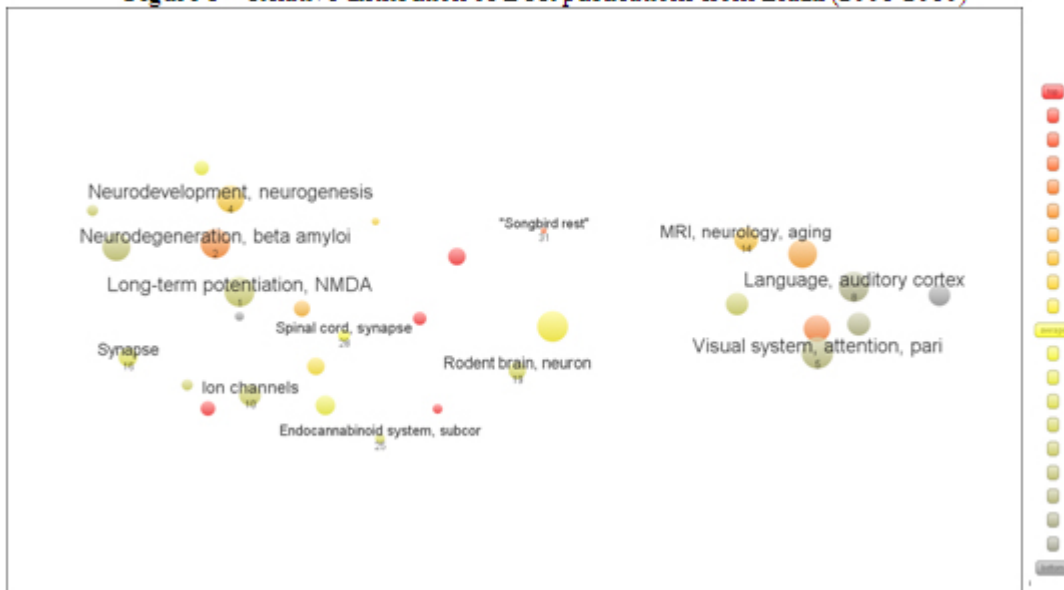


The map shows a clear Dutch preference for Cognitive neuroscience (right-hand side). This preference coincides with a strong growth in the same area world-wide. This explains the increased contribution of the Netherlands in the field overall (from 2% to 4%).

3.2 Case study #2 Brazil in BCR

For Brazil the situation is slightly different. It shows a similar (relative) growth in BCR (Figure 4) as the Netherlands but the focus is not the same.

Figure 8 – Relative distribution of BCR publications from Brazil (2001-2010)



The map reveals that Brazil is almost all over the map but there seems to be not such a clear focus on the **basic neuroscience** area or the **cognitive neuroscience** area. Brazil is almost absent in some of the sub-domains in the **cognitive neuroscience** area but overrepresented in others from the same area. For instance, in the sub-domain we identified with an outstandingly fast growth (**Affective-social; amygdale, autism**) Brazil has a relatively high share. In the *basic neuroscience* area, the sub-domain with a negative growth (**Long term potentiation, NMDA**) Brazil is almost absent. But there is another interesting aspect of the Brazilian focus. In the interface area between basic and cognitive area (the middle part of the map) we found an increased share of Brazil. Although most of the sub-domains in this area are small (relatively small numbers of publications) the Brazilian share is as expected or higher. This interface area does not show the same fast growth as the cognitive neuroscience area but it does grow and experts have indicated this as an important area for the field as a whole.

The most active organizations in the **interface** area are *University São Paulo* (USP), *Catholic University Rio Grande Sul* and *Federal University Rio Grande Sul*. In the **basic neuroscience** are the most active ones are *Federal University Rio de Janeiro* (UFRJ), *USP* and *Federal University Minas Gerais*. In the **cognitive neuroscience** area they are *USP*, *State University Campinas* and *UFRJ*.

4 Conclusions and discussion

The approach to explore a science field and position actors within, using mapping and related techniques appears to be useful to go on step beyond numbers of production as only. With the applied approach we were able to enhance simple figures on activity and growth. Such approaches will become more important as bibliometrics appears to become more attractive to science policy makers and research managers. The threat of 'spreadsheet management' requires more sophisticated methods and tools.

In this study we illustrated the potential of science mapping within this context with a case study in BCR and the position of two countries within. With this approach we could explore the developments in the field in detail and draw general conclusions for the field. We found that in BCR particularly in the area of cognitive neuroscience, there was a sharp increase of activity, as measured by numbers of papers published. Also in the interface area between basic and cognitive neuroscience we detected an increase of activity, albeit not so pronounced.

For the two countries we studied, we were able to position them using the structure and in more detail interpret their de-

velopment. We found a very clear preference of Dutch BCR in the area of cognitive neuroscience and preference for Brazilian BCR in the interface area.

By differentiating activity and impact (not addressed here) using the map structure, we can more specifically identify important actors in the field and in areas within.

Uso de mapas da ciência no contexto de políticas científicas

RESUMO

This paper discuss the complex relations adopted by Semiology and media theory, to propose a televisual communication oriented semiotics, capable to embrace the specificities of this kind of media process. Thus, after to point the inadequacy of the elements enforcement of a immanent Semiology, or applied to cinema, to the television analysis, it endorses the notion of gender promise as a replacement to the notion of contract or pact, as subjacent to the media communication.

PALAVRAS-CHAVE: Mapeamento da ciência. Redes. VOSviewer.

Referências

ECK, N. J. van ; WALTMAN, L. Text mining and visualization using VOSviewer. **ISSI Newsletter**, v. 7, n. 3, p. 50-54, 2011.

_____. VOSviewer: a computer program for bibliometric mapping. In: LARSEN, B. ; LETA, J. (Eds.). INTERNATIONAL CONFERENCE ON SCIENTOMETRICS AND INFORMETRICS 12., July 14-17 2009, Indiana. **Proceedings...** Indiana, 2009. p. 886-897.

NATIONAAL RESEARCH INITIATIEF HERSENEN & COGNITIE. c2012. <<http://www.hersenenecognitie.nl>>

NOYONS, E.C.M. **Bibliometric mapping as a science policy and research management tool**. Leiden: DSWO Press, 1999. Thesis Leiden University.

WALTMAN, L.; ECK, N. J van; NOYONS, E.C.M. **A New methodology for constructing a publication-level classification system of science**. ArXiv:1203.0532, 2012.

_____. A Unified approach to mapping and clustering of bibliometric networks. **Journal of Informetrics**, New York, v. 4, n. 4, p. 629-635, 2010.

Ed Noyons

*Doutor em Estudos Quantitativos da Ciência.
Pesquisador sênior do Centre for Science &
Technology Studies (CWTS), Leiden University.
Netherlands.
E-mail: noyons@cwts.leidenuniv.nl*

Recebido em: 12/09/2012

Aceito em: 19/12/2012