

Science, Technology & Innovation Indicators

Thematic Paper 2

Scholarly publication patterns in the social sciences and humanities and their relationship with research assessment

A publication commissioned by the Ministry of Education,
Culture and Science.

Authors:

Pim den Hertog (Dialogic)

Cor-Jan Jager (Dialogic)

Arthur Vankan (Dialogic)

Robbin te Velde (Dialogic)

Jaap Veldkamp (Dialogic)

Dag W. Aksnes (NIFU)

Gunnar Sivertsen (NIFU)

Thed van Leeuwen (CWTS)

Erik van Wijk (CWTS)

Publication number:

2010.056-1436

Utrecht, the Netherlands

December 2014

For more information about STI² please visit
the website (www.sti2.nl)

This report is electronically available at www.sti2.nl

Graphic Design: Teatske sanne



NIFU

Nordic Institute for Studies in
Innovation, Research and Education



Universiteit Leiden

Science, Technology & Innovation Indicators

Thematic Paper 2

Scholarly publication patterns in the social sciences and humanities and their relationship with research assessment

2014





Table of contents

1. Introduction	7
2. Policy and assessment context in the Netherlands	8
3. A debate on bibliometrics and the SSH	10
4. The heterogeneity of the publication patterns – an example drawn from METIS	10
5. Scholarly peer-reviewed SSH output and its coverage in commercial citation indexing	11
6. Distinct features of publication patterns in the SSH	16
7. Comprehensive coverage using institutional data	18
8. The model proposed to assess research in the social sciences and humanities	19
9. Discussion	21
References	23
Dutch Summary	25



Thematic Paper 2

Scholarly publication patterns in the social sciences and humanities and their relationship with research assessment

1. Introduction

The social sciences, humanities, and law (SSH) are known to have more heterogeneous publication patterns than the sciences, medicine, and technology (STM). On the one hand, original peer-reviewed research is published in a wider range of formats. Book publishing (monographs or articles in edited volumes) may even be more important than journal publishing in some of the disciplines (Sivertsen & Larsen, 2012). On the other hand, scholars in the SSH, more often than their colleagues in the STM, publish directly for a wider audience in the societies and cultures that they relate to in their research (Bentley & Kyvik, 2011). These publications are both popularizations of research in general and professional communication of specific results, e.g. in commissioned reports. Even the peer-reviewed scholarly publications in the SSH may appear in the national language if this is more relevant with regard to content and outreach (Hicks, 2004). In addition, nationally adapted textbooks for students are often preferred over international standard editions. Consequently, SSH scholars often appear as authors of textbooks and other educational material.

As we shall see, the SSH disciplines are heterogeneous themselves in their publication patterns. But generally, you find that those SSH scholars who are not from English speaking countries will be publishing in a minimum of two languages, one of which is their native language and the other the dominant international language of the field (which in certain humanistic disciplines need not be English). Indeed, there is a gradual and stable increase in English language publishing in the SSH, but there are also large differences between the disciplines (van Leeuwen, 2006; Ossenblok et al., 2012; van Leeuwen, 2013), indicating that the bilingual situation will prevail in the SSH due to the societal obligations and wider audiences, as explained above.

In this paper we will describe the written communication cultures in the SSH domains and how they differ from the STM domains, by looking at the type and language of publications, the use of authorship and citations, and the audiences. We will also look at the coverage of the publications in the bibliographic data sources that are normally used for bibliometrics in support

of research assessment. We will discuss to what extent these data sources – which are limited from the point of view of the SSH – can be supplied by institutional data in current research information systems. We describe the current policy context and development of these issues in order to demonstrate which improvements could be most beneficial for the research assessment of the SSH domains.

2. Policy and assessment context in the Netherlands

In the Netherlands, from the early 1990s onwards, research evaluation has been organized initially under the responsibility of the Association of Dutch Universities (VSNU) alone, and from 2003 under the combined responsibility of VSNU with the National Research Council (NWO) and the Royal Netherlands Academy of Arts & Sciences (KNAW). This resulted in the so-called *Standard Evaluation Protocol* (SEP), which describes in detail how research assessment should be organized. This evaluation protocol is periodically revised, and in every cycle of six years, the development of the next evaluation protocol starts immediately. The cycle consists of a midterm review, in the form of an internal self-evaluation study, and an international assessment with an external peer review committee. It is crucial to understand that peer review was and is the main guiding principle for how research assessments are conducted. In the first ten years this system existed, the main responsibility for the organization of the assessments was within the so-called ‘Chambers’ of the VSNU. As these ‘Chambers’ were populated with representatives of the various disciplines accommodated by the universities, they were the best equipped group of dedicated scholars to start thinking about and organizing research assessment. In many of the research assessments taking place in the 1990s, bibliometric analysis accompanied the research assessment processes, although this was not an explicit part of the evaluation protocol. Only in those fields where bibliometrics would make a relevant contribution, was advanced bibliometrics introduced as part of the data available to the peer review committee (van Leeuwen et al., 1996, van Leeuwen et al., 2002). In the 1990s, biomedical research was assessed by KNAW only, without any assistance from VSNU and NWO.

It is important to realize that the evaluation protocol is the formal assessment guide for evaluating academic research in the Netherlands, as well as the institutes of both NWO and KNAW. Furthermore, it is also the informal guide for assessing research that does not come under the supervision of any of these three stakeholder organizations. In the SEP, the focus is on the assessment of research performance at the institutional as well as the research programme/group level; it is not the intention to judge individual researchers. The topics SEP focuses on are scholarly performance, the role of education and teaching in relation to PhDs, and various other topics (over time, we have seen a shift in focus on topics such as interdisciplinarity, valorization, and more recently, societal impact). In 2003, the SEP put the main responsibility for organizing research assessment in the hands of individual university Boards. This created a much more fragmented organization of research assessments, in which the former national perspectives were somewhat lost.

In the evaluation protocol, the focus is on four different indicators of scholarly performance, namely Quality, Productivity, Relevance, and Vitality. In the protocols up till now, the assessment

on these four criteria has been made on a five-point scale. Academic quality is based on the quality of the research group's output: dissertations, academic publications, professional publications (where relevant), patents (where relevant), and other academic products (tests, prototypes, software). Important elements of the assessment are the academic level of the publications, originality, and coherence, while at the same time the contribution to the discipline/field/specialty is considered. The quality of the achievements is also reviewed from the perspective of the group members' standing, such as contributing and collaboration in an international context, editorial board memberships of international journals, international conference contributions, etc.). Productivity is assessed as the total output from the perspective of human resources involvement. For reasons of comparison, the committees are supposed to take into account differences in publication practices among disciplines, and the different position of some scholarly communities in society at large, as can be learnt from their mission statements (in the self-evaluation reports). The element of relevance is considered as the extent to which a unit's research contributes significantly to the development of a field, as well as the decisions made for advances in research. Furthermore, for those units with a strong strategic/applied focus, academic relevance is coupled with societal/technological impact. Just as in the assessment of productivity, we should consider the mission of the units under assessment. Finally, viability is supposed to be assessed from the direction in which the unit under assessment is evolving, as expressed in the topics chosen by the units, and the link between their research agenda and the international context.

In 2003, a major revision of the evaluation protocol, from that moment on called SEP (Standard Evaluation Protocol), resulted in a shift in the VSNU Chambers' responsibility, returning it to the individual university Boards. This change made the research assessment more local, and possibly also less at a 'distance', as the previous construction might have been perceived. Because research assessment had become too much of just an administrative exercise, concerns were growing about the increasing distance between the governance of research within the institutions, and work floor level. Within the social sciences and humanities community, concerns were mounting about the design and application of the evaluation protocol, particularly since more and more bibliometrics were being introduced in research assessment procedures. This led to KNAW setting up an advisory committee consisting of high profile scholars from the social sciences and humanities. The report this committee produced, 'Judging research by its merits' (Council for the Humanities [Raad voor Geesteswetenschappen] and Social Sciences Council [SWR], 2005) clearly described the need for adequate evaluation procedures in the SSH domain. This report instigated various follow-up initiatives in the social sciences, the humanities, and the engineering sciences, on how research assessment could be best organized and designed, and the creation of the types of indicators that would best fit the various communication cultures. Three separate advisory committees prepared reports on how the research assessment procedure would fit these three domains of scholarly activities (KNAW 2011, KNAW, 2011, KNAW, 2013). All three of these advisory reports made the distinction between scientific impact and societal impact. Two of the three reports, namely for the social sciences and the humanities presented a more elaborate model. We use this model in our study to elaborate on research assessments in the social sciences and humanities. Interestingly enough, the new SEP 2015-2021 has integrated this proposed model as the new blueprint on how to assess research in the Netherlands (see paragraph 8).

3. A debate on bibliometrics and the SSH

As described above, the Netherlands has a 20 year tradition of taking care of research quality, in which peer review has played a central role. Bibliometric analyses have accompanied research assessments as a support tool for peer review committees since the 1990s. As the data source was always the Web of Science and citations were the basis for the main indicators, this only concerned the STM disciplines that are well covered in the data sources. Psychology was included as the only discipline in the SSH.

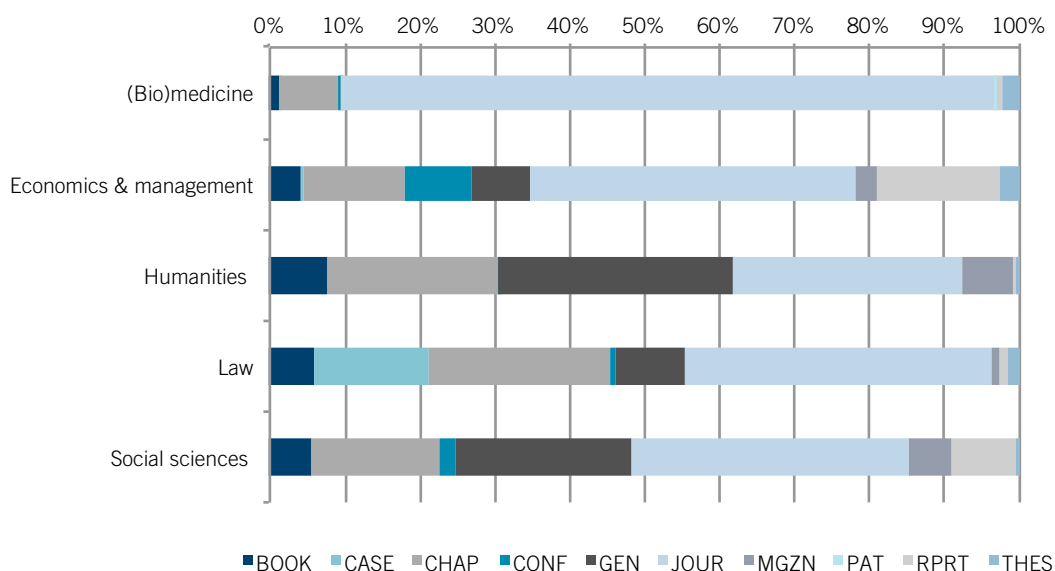
Around the turn of the Millennium, a fierce debate raged in the *Academische Boekengids*, a general periodical for academic communication in the Netherlands (Dehue, 2000, 2001). The direct starting point was the abovementioned assessment of psychology research in the Netherlands, but the publication patterns in the SSH in general soon came under scrutiny. Scholars representing the SSH claimed that an evaluation methodology based on the WoS only was inappropriate in the SSH. Concerned that book publishing was under threat, the scholars strongly opposed the idea that journals, particularly those in English, should become the dominant scholarly communication channel in their fields.

Similar debates were taking place among Dutch speaking scholars in Belgium. In 2003, a financial resource mechanism, the so-called BOF key [*Bijzonder Onderzoeksfonds*] was introduced for Flemish universities, with a bibliometric indicator based on the Web of Science only. SSH scholars protested through their organizations until in 2008, the government provided the legal framework for constructing a supplementary data source, the Flemish Academic Bibliographic Database for the Social Sciences and Humanities [*Vlaams Academisch Bibliografisch Bestand voor de Sociale en Humane Wetenschappen* or *VABB-SHW*], which now provides a comprehensive coverage of non-WoS publications (Engels et al., 2012).

4. The heterogeneity of the publication patterns – an example drawn from METIS

METIS is a system for university publication repositories that has been discussed, but not yet implemented, as a potential comprehensive national database of metadata similar to VABB-SHW or the more comprehensive Current Research Information Systems (CRIS), which cover research activities and results in every domain at the institutional level. Although the coverage of bibliographic references in METIS is not yet comprehensive, it can still provide valuable information on the characteristics of SSH publication patterns, as the following example will show. The example is taken from a university in the Netherlands with a general and comprehensive disciplinary profile. It is therefore possible to compare the SSH patterns with those found in the STM field of biomedicine. Two SSH fields, law and economics & management, have been separated and specified in our analysis. The percentage distribution of publications among ten different publication types is shown in Figure 1.

Figure 1: Composition of the output for domains, one university in METIS, 2004-2009



The legend for Figure 1 reads as: BOOK is books (monographs and edited volumes), CASE is case reports and Law annotations are included under this heading, CHAP is chapters in books, CONF is conference papers, GEN is a container concept for all other types of scholarly activity, JOUR contains peer-reviewed and non peer-reviewed journal publications, MGZN is magazine contributions, PAT stands for patents, RPRT is category reports, and THES is dissertations.

Our results clearly show that the SSH has a broader range of outputs than the STM subfield, where the output is much more concentrated in scientific journals. This highlights the importance of including societal publications and other types of communication with society in the Standard Evaluation Protocols in the Netherlands. For our discussion on *scholarly peer-reviewed publications* in the following section, we should point out that these types of output represent only a part of the total SSH publishing activity.

5. Scholarly peer-reviewed SSH output and its coverage in commercial citation indexing

The heterogeneity of publication patterns, as well as commercial considerations, may explain why the SSH has a more limited coverage than the STM in international and commercial bibliographic data sources such as *Scopus* (Elsevier) and *Web of Science* (WoS, Thomson Reuters). The coverage is limited even if we only compare peer-reviewed publications from original research. Recently, there has been a positive trend towards broader coverage of the SSH in these data sources. The number of SSH journals that are covered has increased, especially in *Scopus*, where the coverage now clearly surpasses WoS. At the same time, Thomson Reuters has supplemented WoS with the *Book Citation Index* (BCI), while Elsevier has included scholarly books and book series from a selected set of publishers in *Scopus*.

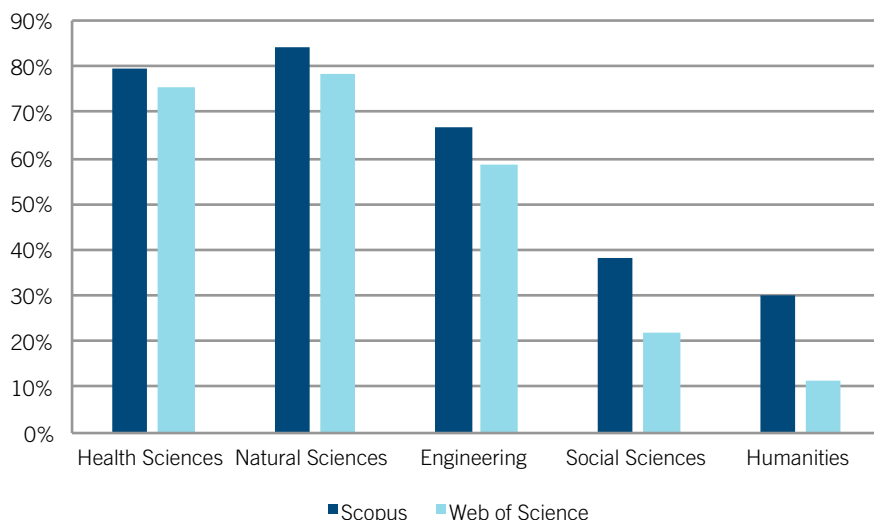
So far, these changes have not altered the picture of limited coverage of the SSH. Below, we will demonstrate the degree of coverage in the three mentioned data sources down to the level of SSH disciplines. As we shall see, there are large variations, indicating that some SSH disciplines will find themselves well represented in the commercial data sources while others will not accept the coverage as a representation of their research efforts. In order to relieve this potential tension within the SSH, it seems important to demonstrate the limitations empirically and discuss them openly, concentrating on how to remedy the situation.

Several studies have demonstrated the limitations of the WoS coverage across SSH disciplines by measuring the proportion of references in indexed documents that relate to other already indexed documents (van Leeuwen & van Raan, 2003; Nederhof, 2006; van Leeuwen, 2013). This is called the *internal coverage* procedure, as it describes the coverage issue within the WoS perspective. Traditionally, the *external coverage* procedure measured the extent of coverage in WoS (or any other electronic database used for citation analysis in research assessment) of individual or departmental publication lists. But recently, it has become possible to compare the total output of countries that have a comprehensive coverage of their scholarly publication output in national databases with international electronic databases such as WoS, Scopus, and others. Given that the SSH disciplines have specific publication patterns that are similar across countries, but not necessarily similar to the neighboring discipline (van Leeuwen, 2006; Ossenblok et al., 2012), we should be able to inform the situation in Netherlands by drawing on data from other countries.

An increasing number of European countries – among them Belgium (Flanders), Croatia, Czech Republic, Denmark, Estonia, Finland, Hungary, Norway, Portugal, Slovenia, Spain and Sweden – has established, or is in the process of establishing, national current research information systems with complete, quality-assured bibliographic metadata for their country's scholarly publication output in the public sector. Such national information systems represent the potential for a more comprehensive coverage of the scholarly literature of the social sciences and humanities (Hicks & Wang, 2011) if connected to a scheme for institutional funding (Hicks, 2012). In 2005, Norway was the first country to establish a national information system with complete quality-assured bibliographic data covering all peer-reviewed scholarly publishing in the entire higher education sector (Schneider, 2009; Sivertsen, 2010). We use a dataset of more than 70,000 scholarly publications from the eight years 2005-2012, 44 per cent of which are in the SSH, to study the SSH publication patterns and their coverage in Scopus and WoS.

Since the Norwegian data include all fields of research, we start by showing an overall comparison of SSH fields with STM fields in Figure 2. It includes both book and journal publishing.

Figure 2: Coverage in Scopus and Web of Science of 70,500 scholarly publications in journals, series, and books from the higher education sector in Norway 2005-2012.



The general picture is that although Scopus has a broader coverage, the two data sources follow the same pattern in representing major fields. We can see from our data that the deficiencies in the SSH are mainly due to incomplete coverage in international journals, limited or no coverage in national disciplinary journals, and very limited coverage in peer-reviewed scholarly books.

In general, articles in journals and series represent only 48% of the publications in the humanities in our data. This result can be compared to the fact that some 45% of all documents in the *Arts & Humanities Citation Index* (A&HCI) are book reviews (Zuccala & van Leeuwen, 2011). Books are also important in the social sciences, where 58% of the publications are in journals or series. In comparison, this proportion is 89% in the health sciences, 88% in the natural sciences, and 71% in engineering. These differences are stable throughout the eight years 2005-2012.

Of all publications in the humanities, 48% are in international languages. The corresponding proportions are 53% in the social sciences, 83% in the health sciences, and 97% in the natural sciences and engineering.

As can be seen in Tables 1 and 2, the percentages are somewhat higher if only journal publications are considered. The use of international language in the SSH has been increasing during the period, but only slightly more than the use of the native language.

Tables 1 and 2 give a more detailed picture of the level of disciplines with regard to publications in *journals*. Note the large disciplinary variations in coverage. As explained in the introduction, we maintain that the differences in coverage of e.g. economics versus law is not only related to Norwegian scholarly publication practices or the coverage of Norway in Scopus and WoS, but is rooted in the differences in publication patterns (and the missions, subject areas, and methods in the field) that are similar across all countries.

Table 1: Scopus and WoS coverage of disciplines in the humanities with special focus on journal publishing. Based on complete data for peer-reviewed publications in the Norwegian HE sector 2005-2012.

Subfield	All publications		Journal publications		
	Total	% in journals	% in int. language	% in WoS	% in Scopus
Classical Studies	259	66 %	48 %	23 %	26 %
Theatre and Drama	129	66 %	61 %	14 %	29 %
Linguistics	1,057	61 %	93 %	39 %	41 %
Ethnology	392	57 %	47 %	12 %	16 %
Literature	764	57 %	31 %	16 %	18 %
Archaeology and Conservation	765	56 %	52 %	26 %	30 %
Slavonic Studies	231	56 %	84 %	17 %	43 %
Architecture and Design	424	54 %	38 %	11 %	24 %
Philosophy and History of Ideas	1,121	54 %	45 %	28 %	33 %
Art History	278	54 %	45 %	21 %	25 %
Musicology	403	50 %	43 %	28 %	26 %
Theology and Religion	2,126	50 %	42 %	16 %	34 %
History	1,645	45 %	44 %	40 %	44 %
Media and Communication	1,073	39 %	73 %	19 %	47 %
Asian and African Studies	237	39 %	99 %	42 %	49 %
Germanic Studies	238	38 %	100 %	39 %	37 %
Romance Studies	304	35 %	100 %	47 %	55 %
Scandinavian Studies	1,777	35 %	17 %	2 %	2 %
English Studies	329	32 %	100 %	39 %	60 %
Total	13,551	49 %	52 %	23 %	32 %

Table 2: Scopus and WoS coverage of disciplines in the social sciences (including law and psychology) with special focus on journal publishing. Based on complete data for peer-reviewed publications in the Norwegian HE sector 2005-2012.

Subfield	All publications		Journal publications		
	Total	% in journals	% in int. language	% in WoS	% in Scopus
Library and Information Science	389	83 %	98 %	56 %	80 %
Psychology	1,940	79 %	79 %	66 %	72 %
Geography	853	78 %	86 %	72 %	78 %
Economics	1,081	75 %	83 %	73 %	77 %
Business & Administration	2,904	63 %	76 %	39 %	57 %
Law	2,108	61 %	31 %	6 %	13 %
Anthropology	597	53 %	65 %	32 %	82 %
Gender Studies	358	48 %	38 %	19 %	37 %
Sociology	1,157	46 %	60 %	40 %	48 %
Political Science	1,655	45 %	76 %	64 %	73 %
Education & Educational Research	4,861	43 %	51 %	22 %	35 %
Total	17,903	58 %	66 %	42 %	54 %

We see that Scopus covers 32%, while Web of Science covers 23% of all peer-reviewed scholarly articles in journals and series in the humanities at Norway's higher education institutions. The corresponding figures for the social sciences (including law and psychology) are 54% in Scopus versus 42% in Web of Science.

The difference between Scopus and WoS is not due to journals being published in Norwegian or Scandinavian languages. With very few exceptions, such journals are not covered in any of the data sources. The number of journals thereby not covered is small, since at the national level, many articles are concentrated in only a few journals (Sivertsen & Larsen, 2012). The difference between the two data sources is that Scopus has a wider coverage of international journals in the SSH. There is no journal covered by WoS in the Norwegian data that is not also covered by Scopus.

While Scopus has a broader coverage of journals, Table 3 shows that the situation is the opposite with regard to scholarly publishing in books. Thomson Reuter's *Book Citation Index* covers 17% of the peer-reviewed monographs and articles in edited volumes in the humanities, while Scopus covers only 5%. The corresponding figures for the social sciences (including law and psychology) are 28% in the Book Citation Index versus 7% in Scopus.

Table 3: Scopus and BCI coverage of disciplines in the humanities with special focus on book publishing. Based on complete data for peer-reviewed publications in the Norwegian HE sector 2005-2012.

Major field	Subfield	Publications in books	BCI	Scopus
Humanities	Classical Studies	278	9 %	2 %
Humanities	Theatre and Drama	163	23 %	6 %
Humanities	Linguistics	129	33 %	19 %
Humanities	Ethnology	175	45 %	9 %
Humanities	Literature	441	23 %	8 %
Humanities	Archaeology and Conservation	80	32 %	12 %
Humanities	Slavonic Studies	799	13 %	3 %
Humanities	Architecture and Design	146	5 %	1 %
Humanities	Philosophy and History of Ideas	112	8 %	0 %
Humanities	Art History	324	41 %	19 %
Humanities	Musicology	266	7 %	1 %
Humanities	Theology and Religion	538	16 %	3 %
Humanities	History	163	21 %	2 %
Humanities	Media and Communication	1,015	4 %	3 %
Humanities	Asian and African Studies	168	26 %	3 %
Humanities	Germanic Studies	93	26 %	1 %
Humanities	Romance Studies	44	0 %	0 %
Humanities	Scandinavian Studies	915	16 %	5 %
Humanities	English Studies	136	24 %	19 %
Humanities	Total	5,977	17 %	5 %
Social Sciences	Library and Information Science	58	14 %	5 %
Social Sciences	Psychology	158	33 %	8 %

Social Sciences	Geography	161	14 %	4 %
Social Sciences	Economics	2,275	8 %	3 %
Social Sciences	Business & Administration	720	20 %	9 %
Social Sciences	Law	228	48 %	22 %
Social Sciences	Anthropology	244	29 %	9 %
Social Sciences	Gender Studies	511	23 %	5 %
Social Sciences	Sociology	773	40 %	7 %
Social Sciences	Political Science	890	35 %	12 %
Social Sciences	Education & Educational Research	345	28 %	8 %
Social Sciences	Total	6,363	22 %	7 %

Interestingly, we find that *Scopus* and Thomson Reuter's *Book Citation Index* has a broader coverage of book publishing in the STM fields that they do in the SSH, where the coverage is quite narrow, even of publications from prestigious international book publishers in these fields. *Scopus* has a very narrow selection of publishers that mainly operate in the STM market. The BCI has a better representation of the SSH, but the coverage of publishers still seems to be in an initial phase, as mainly English language publishers have been selected.

6. Distinct features of publication patterns in the SSH

There are a few other differences between the SSH and the STM domains that deserve mentioning before we discuss the potential solutions for a more comprehensive representation of the social sciences and humanities.

A first important difference relates to the pace of conducting research, and the way scientific knowledge ages in time. While scientific knowledge in the STM domains becomes obsolete within 3 to 4 years, this might take up to ten years or more in the SSH domains, and books published in the SSH domains can influence the field sometimes even for decades (Martin et al., 2012). As seen above, scholarly research in the SSH domains often has a local focus, while the research in the STM domains is more internationally oriented. This is often reflected in the degree of international cooperation found in the STM domains compared to the SSH domains. While it is completely normal to have a high degree of international cooperation in the former, we find a much lower degree of international cooperation in the latter. However, there are wide differences in this respect also within the SSH.

The next difference relates to the authorship and the teams involved in conducting research. Multiple authorship is a common phenomenon in the STM domains, and teams are much larger than in the SSH domains, where scholars often conduct research solely, and publish solely or in small teams (Seglen & Aksnes, 2000; Horta & Lacy, 2011). Based on WoS data, Figures 3a and 3b illustrate the development of authorship over time, for a wide variety of scientific disciplines. In the early 1980s, the natural sciences and biomedicine publications carried around 2-3 authors per publication, the engineering, social, and behavioral sciences around 2 authors per publication, while the humanities and social sciences carried roughly one author per publication. Taking this to the more recent years, we observe that authorship in the natural and

biomedical sciences has on average increased to 5-6 authors per publication, while the humanities and some social sciences still have one author per publication. This leads to an important conclusion: authorship per publication has in general increased over time, for most disciplines, but this increase is so far hardly visible in the humanities.

Figure 3a: Average number of authors per publication in STM disciplines in Web of Science 1981-2011.

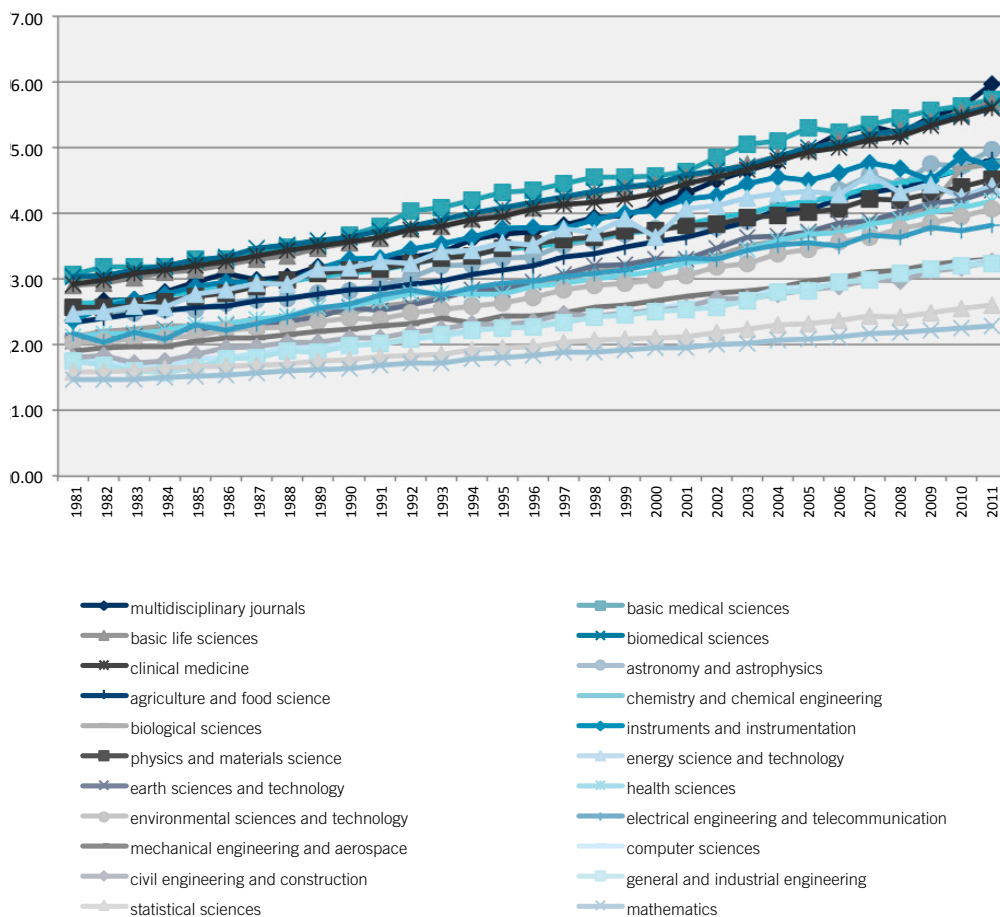
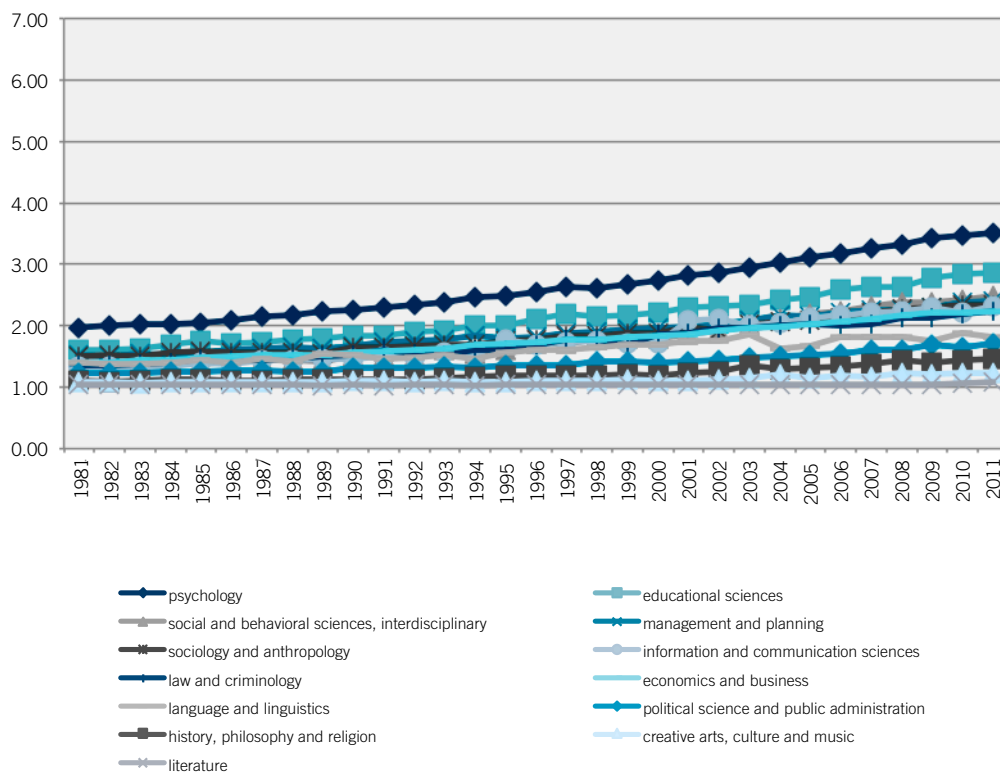


Figure 3b: Average number of authors per publication in SSH disciplines in Web of Science 1981-2011.




7. Comprehensive coverage using institutional data

It is in the interest of research organizations and their funding authorities to attain a complete picture of their research activities and to relieve any tension arising from using specific data sources that only give partial information. The often heard desire to increase the visibility of the SSH is a response to a situation where some areas of research are better represented than others in the most used bibliographic data sources for research assessment purposes.

Up till now, the bibliographic data sources used in library information systems for literature search have not been able to provide comprehensive metadata for the scholarly publications of a specific research organization, e.g. a university. Nor have the digital repositories for institutional archiving with “green” open access to publications yet been able to provide this comprehensiveness. Most promising so far in this respect are the *current research information systems* (CRIS) with a quality assured production of metadata at the institutional (most common) or national level (still very few).

CRIS systems at the institutional level have recently become widespread. They provide not only publication metadata, but also other types of information for the public interface and for management, statistics, and assessment – as a kind of dynamic “annual report” which also serves



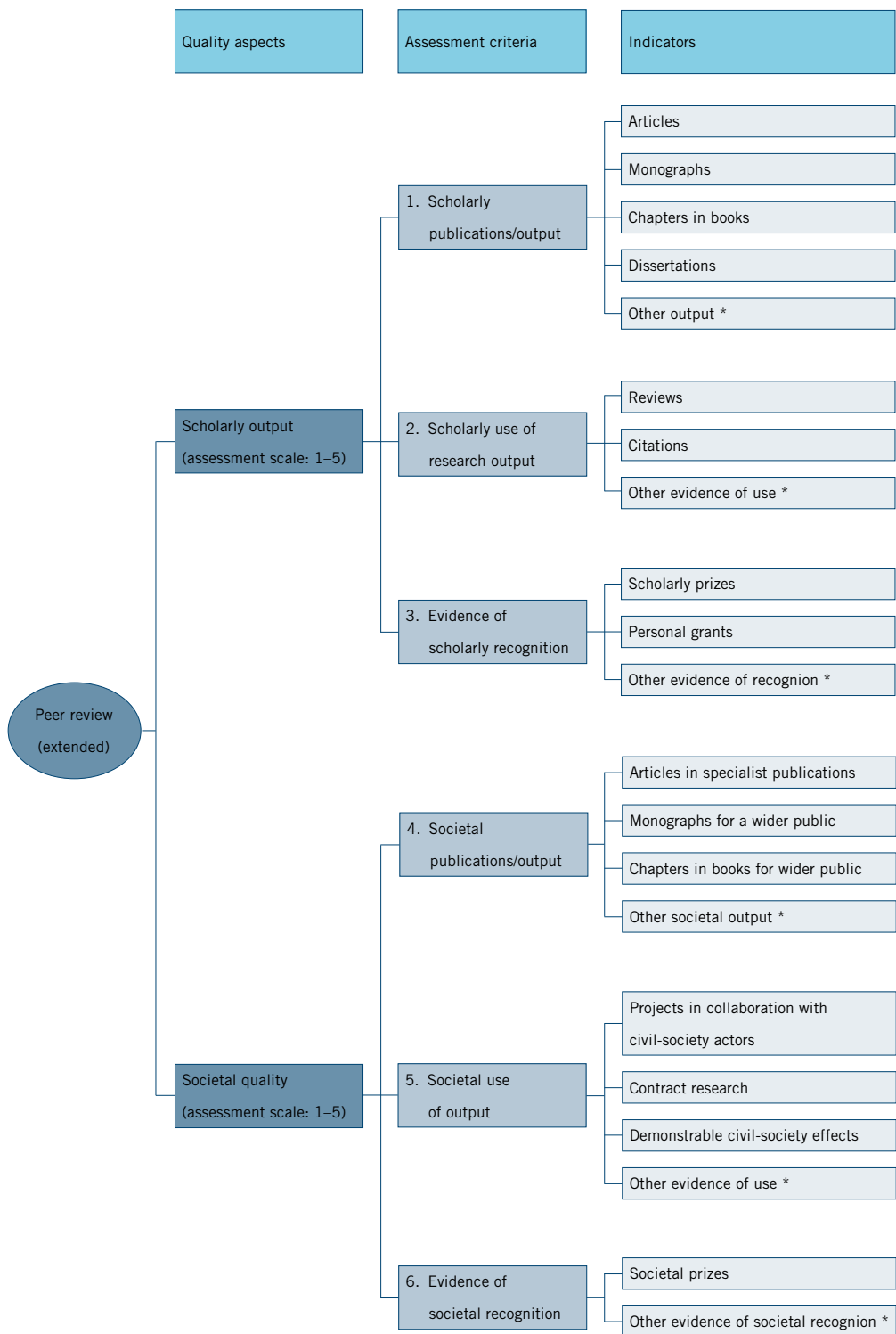
as a communication channel and working space. Institutional CRIS systems come in individual non-commercial solutions as well as through commercial products designed specifically for the purpose. The two leading solutions on the European market have recently been acquired by Elsevier (PURE) and Thomson Reuters (Converis). These solutions have now been integrated with Scopus and the management tool SciVal by Elsevier, and with Web of Science and InCites by Thomson Reuters. The result is that the commercial providers are already working somewhere between limited and comprehensive coverage of the humanities and social sciences.

As mentioned, only a few countries have fully operational and integrated CRIS systems at national level (e.g. the Czech Republic, Estonia, Portugal, and Norway), but several countries have developed, or are in the process of developing, national databases over and above various local CRIS in their higher education sector: Denmark, Finland, Flanders, and Germany. Although these solutions achieve a more comprehensive representation of the SSH publication patterns, the options of citation statistics and international benchmarking seem to be still several steps ahead in this rapid development.

8. The model proposed to assess research in the social sciences and humanities

As peer review is considered the overarching principle in research assessments, its importance forms the starting point for the model proposed by the committees Algra (for the humanities, KNAW, 2011) and Bensing (for the social sciences, KNAW, 2013). From this starting point, scholarly activities are assessed according to two quality aspects, namely the scholarly output and the societal quality. For both of these quality aspects, three quality criteria apply: production, usage, and recognition. Next, for each of these quality criteria, there is a list of indicators. The list of indicators presented has been made short on purpose, it serves as a list of possible examples, and is not intended to be exhaustive. The diagram in Figure 4 shows the schedule proposed in the humanities report.

Figure 4: Diagram of the evaluation schedule for the humanities (KNAW, 2011)



* discipline-specific/context-specific indicators

Both national and international developments have influenced the thinking around the new evaluation protocol in the Netherlands. The declaration made by the community of molecular biologists and cell biologists in San Francisco in 2012 (since then known as the DORA Declaration), criticized the role of metrics and in particular the Journal Impact Factor in research assessment contexts (DORA, 2012). Similarly, the Science in Transition movement in the Netherlands took a critical stance on the existing publication practices and the role of performance based metrics in rewarding scholarly activities (Science in Transition, 2013). These developments, which were taken into consideration for preparing the new SEP launched in 2014, apply to the period 2015-2021, and are largely structured according to the diagram in Figure 4. An additional feature is the narrative, similar to the Research Excellence Framework (REF) in the UK, which also significantly supports assessing the societal impact of research.

Comparing the variety of written communication outputs in the SSH as shown in Figure 1 based on METIS, with the SEP criteria and indicators in Figure 4, we can clearly see how the latest developments of the evaluation protocol in the Netherlands meet the need for appropriate research assessment in the SSH. What is still needed, is a solution for the provision of data for the assessment. It is therefore interesting to note that in other countries, the tradition of recurrent “submissions for evaluation”, which was laborious for the researchers and their institutions, is now being replaced by national or institutional research information systems which run continuously, also for other internal and external purposes than research assessment.

9. Discussion

With a new evaluation protocol design, strongly influenced by the SSH domains themselves in a bottom-up fashion to better fit the various products of scientific communication, the assessment of scholarly activity could be organized in such a way that best serves the diversity and heterogeneity of the SSH. But how do we implement this in practice? Thereby, two main aspects are crucial. Firstly, the collection of comprehensive, verified information and data for the assessment in both dimensions (scientific and societal) as well as on the three criteria (which together form a 2-by-3 matrix); and secondly, the appreciation attributed to these products (in the form of weights, as a result of measurement and comparisons).

Regarding the first aspect, given that Dutch universities have not agreed on METIS as a shared solution, we need to examine their individually chosen alternative systems and compare how these systems might perform according to the requirements of the SEP. Looking at solutions in other countries might also be helpful.

Regarding the second aspect, the appreciation attributed to the products is much more complicated. While the assessment criteria particularly on *Usage* and to a lesser extent *Recognition* are more elaborate for the STM domains in citation indicators, they are difficult to develop in a bibliometric manner for the SSH domains due to problems with data coverage and validity. As we have seen above, the coverage of the SSH in Scopus and Web of Science is still limited. Other sources for citation impact analysis have not yet been fully developed, however, experiments are being carried out using non-source items such as books (Nederhof et al., 2010; Chi,

2013), book citation index systems (see above), or Google Scholar (see Prins et al., 2014). But these methods are far from ready to handle large and consistent data as is possible in the STM domains. On the other hand, the new evaluation protocol allows more scope for other ways of assessing scientific and societal quality or impact than citation indicators. The challenge is staying close to the uniqueness of the SSH domains when communicating with scholarly and societal peer groups.

An attempt to add value to the SSH output is the weighting with peer judgments of the quality of journals into categories or levels. The best known example is the European Reference Index for the Humanities (ERIH), created by the Standing Committee for the Humanities of the European Science Foundation. However, as of 2014, the categorization of the journal lists in ERIH will not be continued. Instead, the lists will now become a dynamic register of approved peer-reviewed journals, also covering the social sciences, under the name ERIH PLUS. If a country or an organization needs journals ranked, they will have to perform this procedure themselves, still relying on the register as an updated representation of well-performing scholarly journals in the SSH.

There are, however, several examples of journal rankings in the SSH that have been established at the level of individual countries. The so-called “Norwegian model” for ranking journals and book publishers on two levels is now used in Denmark, Finland, Norway, Portugal, and at individual universities in Sweden. The ranking is done in each country, but can be compared across countries. This model was recently evaluated after ten years in Norway, and will be continued there. Its main asset is that it stimulates publishing in the major international journals and book publishers (level 2) while at the same time giving just as much weight to national publishing as in less visible channels abroad (the normal level 1, which contains 80% of the publications). Not only English language publication channels are allowed on level 2, but also e.g. Norwegians publishing in the most prominent Spanish language journals and publishers of Spanish philology. Several other countries have similar models for ranking journals, series and publishers, e.g. Croatia, the Czech Republic, Poland, and Slovenia.

So how does the Dutch situation compare to countries where CRIS systems are installed and operational? The Current Research Information System in Norway (CRISTIN) covers and integrates Norway’s entire science system in the public sector, with registration on a national level by all institutions, also outside the higher education sector. The Netherlands does not have such a national system, and attempts to create one have failed. Currently, efforts are being made to emulate the situation in Denmark, where well-defined bibliographic data are harvested annually by the Danish Ministry of Science & Education from the individual CRIS systems at each university (all use a local version of PURE). This annual procedure in Denmark is linked to a bibliometric indicator in the funding formula, a situation that does not exist in the Netherlands, but in several other countries that have integrated local CRIS data at the national level, e.g. the Czech Republic, Finland, and Portugal. The important element for this current discussion is the technical feasibility of harvesting local data for national or shared purposes. In the Netherlands, this purpose would be to provide comparable information for peer review according to the Standard Evaluation Protocol. More concretely, this could lead to the SSH communities taking the following actions based on the new SEP: In the first place, operationalizing the indicators mentioned in the SEP, on the basis of their fit for the various sub-specialties within the SSH communities;

secondly, coming to an agreement on how to register the various products of scientific activities in the light of the indicators agreed from the perspective of a mutual comparability. It is important to remember that the indicators should reflect the variability and heterogeneity of the research conducted and the way this is communicated within the sub-specialties.

References

- Bentley, P., & Kyvik, S. (2011). Academic staff and public communication: a survey of popular science publishing across 13 countries. *Public Understanding of Science*, 21 (1), 48-63.
- Chi, Pei-Shan, 2013: The characteristics of highly cited non-source items in political science. in: Hinze, Sybille / Lottmann, André (Eds.), 2013: *Translational twists and turns: Science as a socio-economic endeavor*. Proceedings of STI 2013 Berlin – 18th International Conference on Science and Technology Indicators. Berlin: iFQ
- Council for the Humanities (Raad voor Geesteswetenschappen) and the Social Sciences Council (SWR) (2005). *Judging research on its merits, An advisory report by the Council for the Humanities and the Social Sciences Council*, Amsterdam.
- Dehue, T. (2000). Alle boeken de wetenschap uit. Om te beginnen die in het Nederlands, *Academische Boekengids*, 23, 10-12.
- Dehue, T. (2001). Pas op je tellen (de onzuiverheden van citatie-telling), *Academische Boekengids*, 29, 13.
- DORA (2012), the San Francisco Declaration on Research Assessment (<http://www.embo.org/news/research-news/research-news-2013-san-francisco-declaration-on-research-assessment>)
- Engels, T.C.E., Ossenblok, T.L.B., & Spruyt, E.H.J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000-2009. *Scientometrics*, 93 (2), 373-390.
- Hicks, D. (2004). The four literatures of social science, contribution to the *Handbook of Quantitative Science and Technology Research*. The use of publication and patent statistics in studies of S&T systems, edited by H.F. Moed, W. Glänzel, and U. Schmoch, Kluwer Academic Publishers, 473-496.
- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41, 251-261.
- Hick, D & Wang, J. (2011). Coverage and Overlap of the New Social Sciences and Humanities Journal Lists. *JASIST* 62 (2), 284-294
- Horta, H., & Lacy, T.A. (2011). How does size matter for science? Exploring the effects of research unit size on academics' scientific productivity and information exchange behaviors. *Science and Public Policy*, 38, 449-460.
- KNAW, Quality assessment in the design and engineering disciplines: a systematic framework, Council for Technical Sciences, Mathematical Sciences and Informatics, Physics and Astronomy and Chemistry (TWINS Council), Second edition, April 2011.
- KNAW, Quality indicators for research in the humanities, Committee on Quality Indicators in the Humanities, May 2011.
- KNAW, Towards a framework for the quality assessment of social science research, Royal

Netherlands Academy of Arts and Sciences “Quality Indicators in the Social Sciences” Committee, March 2013.

- Martin, B.R., Nightingale, P., & Yegros-Yegros A. (2012) Science and technology studies: Exploring the knowledge base, *Research Policy*, 41, 1182-1204.
- Nederhof, A.J. (2006) Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics*, Vol. 66, pages 81-100
- Nederhof A.J., van Leeuwen, T.N. & van Raan, A.F.J. Highly cited non-journal publications in political science, economics and psychology: a first exploration, *Scientometrics*, 2010, 83 (2), 363-374
- Prins A., Costas, R., van Leeuwen, T.N., & Wouters, P.F. Using Google Scholar in research evaluation of social science programs, with a comparison with Web of Science data, STI 2014 Conference, Book of Proceedings, pp. 434-443
- Schneider, J.W. (2009) An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway, *European Political Science*, 8, 364-378.
- Science in Transition (2013), (<http://www.scienceintransition.nl/>)
- Seglen, P.O., & Aksnes, D.W. (2000) Scientific productivity and group size: a bibliometric analysis of Norwegian microbiological research. *Scientometrics*. 49, 125-143.
- Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISSI Newsletter*, 6(1), 22–28.
- Sivertsen, G (2012). The representation of the social sciences and humanities in the Web of Science-a comparison of publication patterns and incentive structures in Flanders and Norway (2005-9) *Research Evaluation*, vol 21 (4), 280-290
- Sivertsen, G. & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential. *Scientometrics* (2012) 91:567–575.
- van Leeuwen, T.N., Rinia, E.J. & van Raan, A.F.J. (1996). Bibliometric profiles of academic physics research in the Netherlands, CWTS-Report to the Dutch Association of Universities (VSNU).
- van Leeuwen, T.N., Visser, M.S., Moed, H.F. & Nederhof, A.J. (2002). The third bibliometric study on chemistry research associated with the Council for Chemical Sciences of the Netherlands Organization for Scientific Research (1991– 2000), CWTS-Report to the Dutch Association of Universities (VSNU) and the Dutch Research Council (NWO).
- van Leeuwen, T.N. & Van Raan, A.F.J. (2003) Publication output and impact of Delft University of Technology within the boundaries of the ISI databases, Report to the Board of Delft University of Technology.
- van Leeuwen, T.N. (2006). The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible, *Scientometrics* 66, 133-154.
- Van Leeuwen, T.N. (2013), Bibliometric research evaluations, Web of Science and the Social Sciences and Humanities: a problematic relationship ? pages 1-18, published online, *Bibliometrie - Praxis & Forschung*, <http://www.bibliometrie-pf.de/article/viewFile/173/218>.
- Zuccala, A. & van Leeuwen, T.N. (2011), Book reviews in humanities research evaluations, *JASIST*, 62, 1979-1991.

Dutch Summary

Gunnar Sivertsen (NIFU,-Norway) en Thed van Leeuwen (CWTS, Universiteit Leiden)

In het tweede thematische paper van deze WTI² cyclus wordt geschetst hoe de sociale – en geesteswetenschappen zich manifesteren qua publicaties, hoe die publicaties ontsloten worden, en hoe die publicatie-output zich verhoudt tot de evaluatie-praktijk, in het bijzonder het nieuwe Standaard Evaluatie Protocol (SEP).

Vanuit bibliometrisch perspectief verschillen de sociale -, geestes- en rechtswetenschappen fundamenteel van de natuur -, levens -, biomedische -, en technische wetenschappen. Dit komt tot uiting in de snelheid waarmee het proces van onderzoek in de sociale -, geestes- en rechtswetenschappen zich voltrekt, de manier waarop de onderzoekers uit de sociale -, geestes- en rechtswetenschappen onderling communiceren, de taal waarin dat gebeurt, en de wijze waarop men samenwerkt. Daarnaast staan de onderzoekers uit de sociale -, geestes- en rechtswetenschappen in een andere relatie tot de samenleving waarin hun onderzoek plaatsvindt, omdat hun onderzoeksvragen vaak direct vanuit een maatschappelijke agenda worden bepaald.

In deze studie laten we zien hoe de publicaties van de onderzoekers in de sociale -, geestes- en rechtswetenschappen worden ontsloten ten behoeve van bibliometrische onderzoek, en de conclusie kan niet anders zijn dat dat daar nog veel aan mankeert. De bestaande bibliometrische analyse is nog onvoldoende uitgerust om de sociale -, geestes- en rechtswetenschappen zodanig in beeld te brengen, dat die in evenwicht is met wetenschappelijke impact meting in andere domeinen. Dit als gevolg van het feit dat een belangrijk deel van de output van de sociale -, geestes- en rechtswetenschappen in lokaal georiënteerde tijdschriften, boeken, en hoofdstukken in boeken niet wordt verwerkt in systemen die gebruikt worden voor bibliometrische analyses.

Met de opkomst van institutionele registratie-systemen is de totale output van de universitaire onderdelen in de sociale -, geestes- en rechtswetenschappen meer en meer beschikbaar gekomen. Dit geeft handvatten om, naast wetenschappelijke impact meting, ook naar maatschappelijke impact meting te gaan kijken. Deze vraag is actueel, omdat er in het recent gelanceerde SEP ruimte is gekomen om ook dat specifieke onderdeel van de wetenschappelijke activiteiten van onderzoekers in de sociale -, geestes -, en rechtswetenschappen te betrekken bij onderzoeksevaluaties.

