

INTERNATIONAL SCIENTIFIC CO-OPERATION AND AWARENESS WITHIN THE EUROPEAN COMMUNITY: PROBLEMS AND PERSPECTIVES

H.F. MOED, R.E. DE BRUIN, A.J. NEDERHOF, R.J.W. TIJSSEN

*Centre for Science and Technology Studies (CWTS/LISBON), University of Leiden,
P.O. Box 9555, 2300 RB Leiden (The Netherlands)*

(Received December 12, 1990)

International scientific co-operation (ISC) and awareness are topics of increasing interest for both scientists and science policy makers. In this paper, we adopt primarily the science policy point of view. After a concise overview of the literature we summarize the main results of the research we conducted. The main outcome with respect to ISC is that it increases. However, large differences exist between countries and between scientific disciplines. ISC and awareness constitute a complex phenomenon, affected by several factors, science-internal, as well as external. In the paper several techniques are described, amongst which those that can visualize ISC relations through analytical maps. An important aspect of our research methodology is the combination of various quantitative, bibliometric analyses and qualitative research on the structure of science and the relations between science and society. Finally, we sketch perspectives for future research.

1. Introduction

The subject of this paper is the phenomenon of integration of activities among scientists from different countries. International integration is conceived as the participation of scientists in international communication networks on research questions, methods, and findings. This topic is increasingly important in science, for science policy makers as well as for scientists in the field. In this paper, we focus on two main aspects: international scientific co-operation (ISC) and international awareness of research activities and findings. By ISC we understand co-operation between scientists from different countries, to be operationalized by trans-national publications, i.e. papers co-authored by scientists from at least two different countries. We consider international awareness as reflected in citations found in scientific literature.

The purpose of this paper is to sketch a number of theoretical issues related to the phenomena of international scientific co-operation and awareness, and to

indicate ways of studying these issues more closely. We will give a concise overview of the main literature in the field and propose an agenda for future research.

International scientific co-operation and awareness may be approached from several sides. In this paper, the main view point is that of science policy. For science policy makers, it is extremely important to have insight into patterns of integration and into the factors influencing these patterns. Such a knowledge may shed light upon the possibilities and limitations of research planning. As far as our research method is concerned, we use bibliometric (i.e. quantitative bibliographic) analyses, combined with a literature-based study on networks of scientific communication and international relations in general.

The paper is structured as follows. First, the problems of international scientific co-operation and awareness are sketched. Here a concise overview of the principal literature on the subject is given. Based on these results, the present knowledge on communication patterns is summarized. In the research questions, especially the possible factors behind constitute an important aspect. Making clear such factors may provide crucial information for science policy. In Section 3, we will deal with the techniques applied, particularly those on 'maps', and give an overview of the research our group (CWTS) performed on the subject until now: bibliometric studies on international co-operation and awareness at the request of panels evaluating EC programmes on agricultural and on medical and health research; a study on ISC in different fields of the natural and life sciences, the social sciences, and the humanities from a global perspective for Dutch policy purposes; the influence of external factors in long-term perspective; and finally a comparison of international literature databases in the three main fields. In Section 4, we will summarize the main problems, sketch the perspectives, and draw an agenda for future research.

2. Perspectives and problems of international scientific co-operation and awareness

There is an increasing interest by science policy makers in ISC. For instance, the Netherlands Board for Science Policy (RAWB), from 1 January 1991 onwards Advisory Council for Science and Technology (AWT), presents quantitative data on this phenomenon in its Science and Technology Indicator Reports. The Dutch Minister of Science and Education plans to collect quantitative, bibliometric data on ISC, and to provide the results with interpretations in terms of policy relevant factors. In several reports, the Ministry points out the importance of internationalization of scientific research. The German Wissenschaftsrat and the National Science

Foundation in the USA are further examples of science policy boards interested in the phenomenon of ISC. The US National Research Council initiated already in 1983 an examination of the state of scientific co-operation between the USA and other industrialized countries.¹

Also for international organizations, ISC is important. These organizations try to foster co-operation among their member states in order to give the internal cohesion also a dimension in science. This is evidently so for the European Community (EC). The Commission of the European Communities (CEC) initiated multi-annual research programmes, for instance in agricultural research and in medical and health research. Fostering co-operation between EC member states is one of the explicit objectives of these programmes.² Establishing the extent of increase in ISC, is an indicator of the success of the programmes.³ We performed bibliometric analyses in order to provide two panels evaluating EC research programmes with relevant data.^{4,5} These analyses will be discussed in Section 3.2.

International co-operation among scientists takes on in several forms and will result in different scientific products. One of the possibilities is a common scientific publication (e.g., in books or in journals). However, not all collaborations lead to common publications. For instance, two research groups may co-operate by 'dividing' the activities in the sense that one group focuses on some part of the research and the second group on another part, each publishing on its own part of the common project. Another possibility is that a member from one research group, working for a certain period in another group, publishes with authors from this group, but does not give his 'original' address.⁴ For bibliometric analyses it is necessary to restrict the research to publications in the serial literature. ISC is measured by trans-national papers, i.e. publications with authors from at least two different countries, as indicated by the address.

In the literature on bibliometrics and research evaluation, several authors drew attention to the phenomenon of ISC. Pioneering work has been performed by Computer Horizons Inc. (CHI) in the 1970s.⁶ Essentially, CHI Research collected data on the number of international co-authorships, based on publication data from the *Science Citation Index (SCI)*, produced by the Institute for Scientific information (ISI) at Philadelphia (USA). An international co-authorship is defined as a co-authorship between authors from different countries, so basically the same as what we defined as a trans-national publication. The data collected by CHI were mainly used by national research councils such as the National Science Foundation in the USA and the RAWB in the Netherlands.

A study by *Lomnitz, Rees, and Cameo*⁷ analyses the rise of both national and international scientific groups with common research interests, crossing institutional and national borders. The use of international co-authorship as a science indicator, we find in a study by *Lewisson and Cunningham*.⁸ They used bibliometric data to evaluate research programmes of the CEC. A different approach of the issue is taken by *Schott*.⁹ He performed a study on international influence in science, analyzing the phenomenon as a global scientific network in which countries influence one another. Like *Frame and Carpenter*⁶, he asserts that this influence depends partly on political, cultural, and social factors. In our research, such external factors play an important role as well. Section 3.4 of this paper will deal with it. According to *Schott*, influence can be indicated by bibliographic references in the serial literature. However, the question whether influence can be measured by citations, is subject to debate. *MacRoberts and MacRoberts*¹⁰ claim that only a small part of the actual influence on a publication appears from its references. *Moed and De Bruin*¹¹ hypothesize that awareness rather than influence is monitored by citations. In their view, citations reflect the author's perception of the cognitive and social background in which he locates his own research activities. Others suggest that a citation reflects mainly the usefulness of the cited paper in the context of the citing paper (e.g., *Garfield*¹²).

A serious problem in studies on international scientific co-operation and awareness is the adequacy of the literature databases from which publication and citation data are extracted. The *SCI* offers the possibility to select trans-national papers easily, because the institutional affiliations of all contributing authors are processed. These affiliations can be found in the corporate source fields of publication data. Moreover, these corporate source addresses are unified, though to a limited extent.¹³ The *SCI* extends over virtually all fields of the natural and life sciences, covering about 3,500 journals. Non-journal publications are not processed. To compensate for this, other databases, such as *Excerpta Medica* for the biomedical sciences, may be used. In such specialized databases, subdivisions in scientific fields are often easier to make. However, the *SCI* is the only multi-disciplinary database processing the addresses of all contributing authors, which makes it a powerful tool for the assessment of international scientific co-operation and awareness. Moreover, the *SCI* has a monopoly on citation data.

An additional problem of the *SCI* is an orientation ('bias') on publications from Anglo-Saxon countries, especially the USA. The result is that other regions are underrepresented. This is particularly true for the Third World.^{14,15} However, also for European countries there are problems. In a case study on tropical soil science,

Arvanitis and *Chatelin*¹⁴ concluded that French authors, at least in this particular sub-field, published mainly articles (written in French) in journals from French-speaking countries, which are not processed for the *SCI*. On the other hand, *Carpenter* and *Narin*¹⁶ concluded that the *SCI* coverage is excellent with respect to the core literature in the physical and biological sciences, at least for English-speaking countries. In fact, they note an Anglo-Saxon orientation as well. Our impression is that the bias problem even becomes stronger in recent years, at least in agricultural research.⁴

Nevertheless, the problems with the coverage are not that bad, that they outweigh the unique possibilities of the *SCI* in the assessment of ISC and awareness. In the *Social Sciences Citation Index (SSCI)* another ISI database, the Anglo-Saxon 'bias' is stronger: 80% of the publications originates from English-speaking countries, while 60% originates from the USA alone.¹⁷

3. Research activities at CWTS with respect to international scientific co-operation and awareness

In this section, we summarize the main results of research performed by CWTS with respect to international scientific co-operation and awareness. These activities cover a wide range of aspects. First, we will deal shortly with data collection. Next, we will discuss the research results of the bibliometric analysis of EC programmes with respect to agricultural research and medical and health research. The third subject of this section is ISC in global perspective, a study performed for Dutch policy purposes. Section 3.4 will deal with external factors affecting ISC, particularly in long-term perspective. Next, data analytical techniques will be discussed. Finally, in Section 3.6, a comparison will be made between the *SCI* and the *SSCI*.

3.1 On-line data collection

Data for the assessment of international co-operation and awareness are derived from the on-line versions of the ISI databases: *SCI*, *SSCI*, and *A & HCI (Arts & Humanities Citation Index)* through the host computers DIMDI and Dialog. We applied basically two methods: determining numbers on-line and downloading publication data to analyze these off-line. To determine the number of papers one country or set of countries, say the EC, co-publishes with other countries, two sets of publications are formed: one comprises all publications of EC countries, and the

other set consists of all publications produced elsewhere in the world. The overlap between both sets represents the total number of co-publications between EC and at least one other non-EC country. This method is free from double counts: even if twelve EC authors from various countries co-publish a single paper with authors from eight non-EC countries, the paper is counted as only one instance of EC- non-EC co-publication. This method can be applied to any delimitation of scientific (sub)fields whatsoever. One option is the classification of journals, as arranged by ISI, into sections of scientific sub-fields. Other delimitations can be devised, based on journals, or on other characteristics of papers, such as title-words or the co-citation cluster(s), if any, to which a paper is linked. In this way, large data matrices can be build, with in the rows and columns specific countries or regions, and in the cells the number of co-publications between each pair of countries/regions.

The other approach is that of downloading publication data according to the method developed by *Moed*.^{18,19} Downloading means the storage of information sent by the host to a personal computer (used as interface between the host computer and the IBM main-frame computer of the Leiden University Computer Centre) on harddisk. The publication data collected relate to authors, journal and article titles, article type, and corporate source addresses. The corporate source fields contain data on the institutional and geographical affiliations of the authors and are crucial in determining ISC. For awareness we also collected the references. After transferring the data from the PC to the main-frame, we processed these with help of 'Statistical Analysis System' (SAS), a software package developed by SAS Institute Inc., Cary N.C. (USA). With this package a database can be created for an off-line analysis of publication data. To establish patterns in international scientific cooperation and awareness, country to country matrices were build with respect to corporate source data and references.

3.2 Bibliometric studies on Agricultural Research and Medical Health Research within the European Community

In 1989 and 1990, CWTS performed two bibliometric studies at the request of special panels, set up by the CEC to evaluate its programmes on agricultural and on medical and health research*. In the study on agricultural research^{4,11}), international scientific co-operation and awareness constitute the central issues. The bibliometric study for the evaluation of three multi-annual research programmes, running from

* This work has been sponsored by the Commission of the European Communities.

1975-1978, 1979-1983, and 1984-1988 respectively, focused on three specific questions:

- (i) Has there been an increase in the proportion of scientific papers, in the field covered by the Community's agricultural research programmes, that are co-authored by scientists from two or more member states over the last decade?
- (ii) If so, how does this compare with the proportion of these papers co-authored by a Community scientist and one (or more) from a non-EC country?
- (iii) Do these papers by Community scientists now cite work from other member states more frequently (as a proportion of the total number of citations), and if so, how does this compare with their citations to non-EC authors over the same period?

Data were collected from SCISEARCH (the on-line version of the SCI) with the techniques described in Section 3.1. First, we created a data set of publications with respect to agricultural research, written by scientists from EC countries. We defined the field agriculture by searching SCISEARCH journals with titles containing significant word-stems, taking into account that titles are possibly in different languages. So we used word-stems such as 'agr', 'food', 'milk', and 'lait'. Articles from this set, containing addresses from one of the EC member states, were selected, and from these publications all information available in SCISEARCH was collected for the years 1979, 1982, 1985, and 1988. Especially information on contributing authors and their institutional and geographic affiliations, as well as on the papers listed in the reference lists are crucial for our research questions.

Since the bibliometric data relate to the period from 1979 onwards, and the research programmes to be evaluated even from 1975 onwards, we decided not to include in the data set the three countries entering the EC during the 1980s: Greece, Portugal, and Spain. Their inclusion would disturb a trend analysis. Consequently, member states in our research are: Belgium, Denmark, the Federal Republic of Germany, France, the United Kingdom, Ireland, Italy, Luxembourg, and the Netherlands. Since we did not find any articles from Luxembourg, this country played no role in our analyses.

The total number of publications from the eight 'old' member states fluctuate between 2500 and 2900 per year. Variations are mainly due to changes in the coverage of the SCI. The number of papers differs considerably per member state. The United Kingdom and the Federal Republic of Germany have the lion's share in

the whole data set, with 30 to 40% each. The share of France is about 10%. The number of publications from all three countries fluctuate. The Netherlands shows an increase from 6 to 11% (160 to 300 publications). The other four countries have only small numbers of papers. The *SCI* coverage seems to an important factor affecting the division between the member states.

As indicated above, we measure ISC by trans-national publications. By selecting all such trans-national papers from the EC data set and next all papers co-authored by scientists from two or more member states, we could establish developments in numbers. Table 1 and 2 show the results.

Table 1
Number and percentage of trans-national publications (relative to the total number of publications)

	Years							
	1979		1982		1985		1988	
	N	%	N	%	N	%	N	%
all	139	5.0	144	5.7	227	7.8	329	12.5
EC-EC	17	0.6	20	0.8	33	1.0	48	1.8

Table 2
Share of EC-EC and EC-non EC co-authorships in trans-national publications

	Years			
	1979	1982	1985	1988
% EC-EC	12.2	13.8	14.3	14.3
% EC-non EC	87.8	86.2	85.7	85.7

From these tables results that there was only a slight increase in the share of the intra-EC collaborations in the total number of trans-national publications by EC scientists. The conclusion seems to be that the multi-annual research programmes of the CEC did not result in an increase of intra-EC trans-national papers, significantly faster than the general increase in trans-national publications. However, this conclusion is restricted to several conditions. First, the coverage of the database may have influenced the results. Next, a considerable time delay exists between the moment a co-operation project starts and the moment possible joint papers become visible in the database. Consequently, the effects of the programme running from

1984-1988 could not be examined completely, and it was precisely this programme which emphasized co-operation between EC member states.² Finally, as indicated in Section 2, not all collaborations lead to common publications.

We also investigated awareness of research in other countries, as measured by citations. About half of the citations is to publications from the countries of the authors themselves. This score is partly caused by self citations, i.e. citations by authors to their own previous work. The share of intra-EC citations in the other 50% fluctuates around 27%. However, citations to the countries entering the EC during the 1980s (Greece, Portugal, and Spain) tend to increase. With respect to relations within the EC, we see basically the same pattern as in ISC. Interesting is the trend of citations to US publications. The percentage decreases during the period 1979-1988 from 59.7 to 50.4.¹¹ On the other hand, there is an increase in citations to Switzerland, Sweden, Australia, and India. These outcomes suggest a shift in the international network of awareness. The position of the USA becomes less central, speaking in terms of a centre and periphery model.⁹ However, this shift does not cause a significantly denser network within the EC.

In the evaluation of the programme on medical and health research (MHR), citations were not included. ISC was investigated in three sub-fields covered by the EC programme (aging, hearing impaired and heart diseases), as well as in three 'control' sub-fields (pediatrics, kidney diseases, and radio-therapy). In the latter fields, we found that co-publications between authors from different EC member states increased at a lower rate than co-authorships between scientists from the EC with colleagues from countries outside the Community, particularly in 1989. For the sub-fields, covered by the MHR programme, the number of intra-EC co-authorships and extra-EC co-authorships increased at approximately the same rate. These results suggest that the MHR programme rather avoided the intra-EC share in the total number of trans-national publications from decreasing than that it caused an increase. However, the same conditions restricting the conclusion on agricultural research, are valid here.

3.3 International scientific co-operation from a global perspective: data from the SCI, the SSCI, and the A & HCI

Four of our studies commissioned by the Ministry of Education and Science of the Netherlands, and the Netherlands Advisory Council for Science Policy (RAWB) have addressed global trans-national scientific publications as occurring in major

science fields, and more recently, scientific and scholarly sub-fields, in the natural and life sciences, the social and behavioural sciences, and the arts and humanities. These studies used the *SCI*, the *SSCI*, and more recently, also the *A & HCI*.

Moed²⁰ tapped the *SCI* for the years 1984-1987, using a field classification similar to that used by CHI, eliminating sections referring to the social sciences and the section 'multi-disciplinary sciences'. By linking ISI sub-fields, he formed eight major fields. However, as some journals are classified by ISI in more than one sub-field, articles are double-counted if they are classified in different major fields. Also, publications were not "fractioned" across countries or regions, which meant that one article co-published by, for example, three countries, would add one full article to the productivity counts of each country. Publications were limited to articles, notes, letters and reviews (not book reviews), and to those incorporating references. This last limitation eliminates differences between the on-line and printed versions of the *SCI*. The on-line version of the *SCI* processes also about 700 journals listed in the Currents Contents journals, which lack most of the information characterizing normal *SCI* coverage, such as addresses and references.

A similar study was conducted for the *SSCI* for 1980-1987¹⁷, but without distinction of sub-fields. Recently, more extensive and sophisticated studies were conducted. The first study focused on the *SCI*, using ISI sub-field-classifications. Using Boolean set logic, Nederhof, Moed, and Ramaekers²¹ eliminated double counting of articles due to multiple classification of journals. A new approach was used for the field "multi-disciplinary sciences", containing important journals such as *Nature* and *Science*. Based on the addresses of authors, articles were classified in nine major fields. For the period of the study, 1984-1989, for each of the six countries studied (USA, UK, Netherlands, FRG, Japan, and France) at least 89% of the articles could be allocated. Furthermore, for each of the nine fields and 29 sub-fields studied, trans-national publication rates between countries were computed. In some fields, trans-national publications are becoming dominant. In the field of "astronomy and astrophysics" for instance, nearly two-thirds of the papers of the Netherlands were co-published with at least one other country, while these percentages varied between 49% and 56% (in 1988-1989) for FRG, France, and the UK. Based on samples of articles, various fractionation rules were tested. Each fractionation rule was based on co-publication of countries. Thus, independent from the respective number of co-authors from co-publishing countries, articles with authors from two countries were counted in principle as 0.5 for each country. The Poisson distribution fitted only well for Japan in a field with relatively little co-publication. Even in fields

with high trans-national publication rates, usually roughly 80% of the cases concerns just two countries. For countries with less than 35% trans-national publications in a (sub)field, a fraction of about 0.46 was used, while otherwise about 0.45 was applied. Recently, this method has been further improved by us on a model of trans-national publication.

The second recent study involved the social and behavioural sciences, and the humanities as monitored through the *SSCI* and the *A & HCI*. This study was even more elaborate than that by *Nederhof* et al.²¹, as both indexes were coupled, and journals were recategorized by hand.²² Co-publication is, on average, less frequent in the social and behavioural sciences than in the natural and life sciences, although the level among the former is comparable to that among the latter in the less transnational fields. In the humanities, trans-national publication is rare and occurs only in 0.7% of all articles, letters, notes and review articles. Nevertheless, in a field like linguistics, trans-national publication rates are comparable to those found in the social and behavioural sciences. For a large part, the lack of big science and of trans-national facilities probably causes the lower rate of trans-national publication in the social and behavioural sciences, and the humanities. In addition, a more local orientation in some fields, a greater sensitivity of subject matter to language and cultural differences, and a different view on scholarship may be of some importance.

To illustrate our method, we offer some results from the *SCI* study for the Netherlands.²¹ Between 1984-1985 and 1988-1989, trans-national publication increased considerably in the following main fields for the Netherlands: engineering (from 15% to 21%), environmental sciences (from 20% to 26%), mathematical and computer sciences (from 23% to 28%), biology (from 19% to 26%), agriculture and veterinary research (from 11% to 17%), and general medicine (from 14% to 18%). In none of other main fields, Dutch co-publication decreased. In the natural sciences and engineering, co-publication with the USA generally declines with 5 to 7%. This trend is reversed only for agriculture and veterinary research. In several fields, co-publication with Eastern Europe is important or is increasing. In engineering, trans-national-publications with Eastern Europe increase from 4% to 8%. In chemistry and in mathematical and computer sciences, Eastern Europe (6% to 9%) is of some importance. In agriculture and veterinary research, the importance of Eastern Europe (0% to 7%) also increases. In each of nine major fields comprising the *SCI*, the rate of trans-national publication increased between 1984-1989. In the social and behavioural sciences, trans-national publication increased in psychology, economy,

sociology, and management studies, stayed about level in law and political science, and decreased in the field of education in the same period.

Finally, intra-EC trans-national co-authorships were monitored in nine main fields for eight EC countries (the Netherlands, Belgium, Denmark, FRG, UK, France, Italy, and Spain). We discuss briefly some of the results (see Table 3). The Dutch share in trans-national intra-EC publications varied between about 14% (physics) and about 32% (astronomy). In engineering, the intra-EC trans-national publication rate nearly doubled, while it increased significantly in most other fields. For environmental sciences, the number of intra-EC trans-national publication pairs declined with 16%. In general, however, trans-national publication rates increased in this field. Apparently, trans-national publication increased in environmental sciences mainly with non-EC countries, at the expense of intra-EC publication.

Table 3
Share of the Netherlands in intra-EC trans-national publications in four fields

Field	Share of the Netherlands		Increase intra-EC trans-national publication pairs 1984-85 - 1988-89
	1984/1985	1988/1989	
Engineering	22%	19%	98%
Physics	14%	16%	53%
Astronomy	32%	29%	37%
Environmental Sciences	17%	20%	-16%

3.4 *The influence of external factors in long-term perspective upon international scientific co-operation*

From our bibliometric studies on the evaluation of the agricultural and MHR programmes of the CEC* we concluded, that the effects upon furthering ISC seem to be rather limited. In our view, many more factors affect this process, such as language, culture, distance, and political and economic factors. *Frame* and *Carpenter*⁶ and *Schott*⁹ already mentioned such factors. We decided to investigate these factors more closely with an emphasis on long-term perspectives.

* This work has been sponsored by the Commission of the European Communities.

Striking examples of such historical factors are colonial ties. In several bibliometric studies we established for instance strong co-operation links between France and many of its former (and present) colonies. In the study on agricultural research, mentioned in Section 3.2, more than half of all French trans-national publications with the Third World are co-published with scientists working in (former) colonies.²³ After independence, most of these countries retained intense relations with the mother country. France still interferes in its former overseas empire, especially in Africa.^{24,25} For the United Kingdom a comparable situation exists. After World War II, the British Empire was transformed into the Commonwealth of Nations and after independence most ex-colonies remained member states of this organization. With these countries many scientific collaborations exist.

An example of the effects of long-term cultural relations (and linguistic affinity) as well as short distance, is the relatively strong co-operation of Denmark with the other Scandinavian countries. For Austria and the Federal Republic of Germany the same pattern exists. In disciplines where publishing in the mother language is still rather usual, such as agricultural research, sharing the same language, as Germany and Austria do, is a factor stimulating co-publishing. In general, neighbouring countries relatively often have collaborations.

The influence of ideology and a political system on science is evident for Eastern Europe during the period 1945-1989. Also on ISC these factors probably have a considerable effect. From preliminary studies we performed, we discovered the existence of a dense network of collaborations among the Warsaw Pact countries and their allies in the Third World, with the USSR as leading nation.

Our investigations on long-term factors affecting ISC are based on bibliometric data and on an elaborate study of the background. By background we understand the structure of science in the countries involved as well as political, social, and economic developments possibly influencing international co-operation. This study is based on literature from different disciplines, such as political science, sociology, and history. This aspect distinguishes our study from that by, for instance, *Frame and Carpenter*⁶, who describe external factors in rather general terms, without referring to specific literature. The difference with studies on ISC by for instance *Lubrano*²⁶ is the extensive use of bibliometric data.

Knowledge of long-term influences on ISC is very important for science policy. When such factors deeply affect ISC patterns, this will hamper the impact on ISC of

policy measures. Charting these factors will elucidate the limitations of such measures.

3.5 Visualizing ISC relations: Analytical maps

The availability of tables of numerical data on numbers of ISC papers will – quite naturally – lead to questions such as: which scientific fields, or which countries, have relatively strong ISC relations given world-wide ISC activity, or which countries show a distribution of trans-national papers that is deviant from the general pattern? Obtaining answers to such questions may necessitate simultaneous examination of the whole of relations displayed within the table. However, if the table is large and/or consists of a complex structure of data, such a visual exercise is beyond human capability. Therefore, a certain degree of data reduction is necessary. Instead of focusing the attention on a limited number of items in the table, or using summary statistics, one might consider the use of a graphical aid (a 'map') to visualize the whole relational structure underlying the data. A map offers an easy visual assimilation of the data and reduces the possibility of overlooking (obscured) features in the structure.

Such maps are mostly based on data analytical methods. There are several types of maps at ones disposal, depending on the type of ISC data and the type of questions one would like to answer. As an example of work performed in our group, we present a map derived with a method from the family of multidimensional scaling methods, known under the name Correspondence Analysis (CA). This method can yield an overall view of the structure of ISC interrelationships as expressed in numerical data in any square of rectangular (complete or incomplete) array (the interested reader is referred to *Greenacre*²⁷ for an excellent textbook on CA).

CA-generated maps depict the significant deviations from an 'average' ISC structure found in the data. CA assigns a co-ordinate to each row and column of the array for each dimension of the spatial display. The relations between the row and column elements are then depicted as a configuration of points – mostly in a two-dimensional space. The distance between rows and columns is dependent upon the strength of their relationship, as expressed by the value in the cell of the array on the intersection of the respective row and column – a relatively small distance in the map indicates a relatively high ISC level between the respective row and column element. The accuracy of CA maps depends on the user-specified dimensionality of the spatial configuration and the complexity of the data. The dimensions in the CA

map have no special significance – they are geometrical constructs determined by the data itself.

In Figure 1, CA is used to visualize the position of nine Western European countries, based on their distribution of ISC papers across a set of 18 countries/regions spanning the entire globe (this includes the same set of nine countries). The data were retrieved from the internationally co-authored research literature in the period 1984-1987 as found in the database SCISEARCH, aggregated over nine large scientific areas. The nine countries are thus portrayed twice in the map: (1) as one of the nine row elements – indicated by bold-printed labels, and (2) as one of the 18 column elements – indicated by plain-printed labels. CA places a row element amidst of the countries with which it shares the largest degree of ISC in terms of the relative number of trans-national publications. In this particular case, it can result in a configuration in which a row element and column element representing the same country are found in the same area of the map, but for different reasons (e.g., Belgium).

Looking at the nine countries as a row element, there is a certain similarity with the geographic location of these nations – hence, the map shows that there exists a stronger tendency to ISC amongst those of these nine nations that share a border. Furthermore, it is, for example, shown that the UK has a somewhat deviant ISC distribution; this country has a relatively strong ISC relations with nations in Asia – Japan excluded, Australia and Pacific, and Canada. Another example: France has relatively strong ISC linkages with Belgium, Central and South America, and Africa. These are findings that may not have been so apparent when eye-balling the table itself.

3.6 Differences between the SCI, the SSCI and the A & HCI

The *SCI*, *SSCI*, and *A & HCI* differ greatly in size: the *SCI* covers yearly nearly ten times as many articles, notes, letters, and reviews as the *SSCI*, while the *SSCI* is roughly double the size of the *A & HCI* in this respect. Thus, due to mere size, trends in the *SCI* are statistically more reliable than those in the *SSCI* and the *A & HCI*, especially when the differences in trans-national publication noted in Section 3.2 are also taken into account.

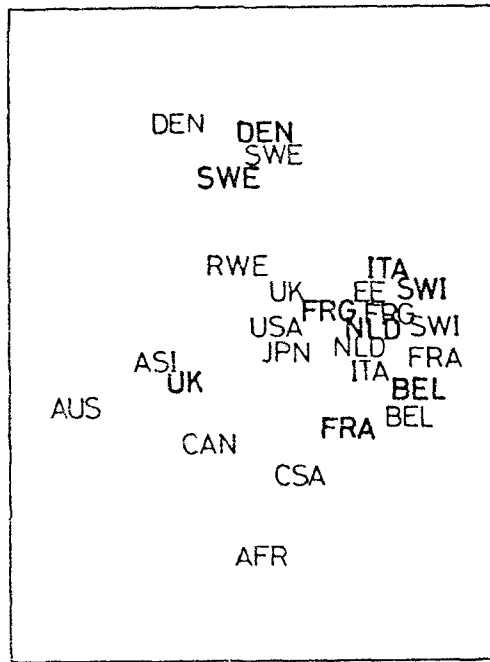


Fig. 1. Correspondence Analysis map of the ISC structure.
 Bold label = selected Western European countries;
 plain label = all countries/geographical regions

Legend:

- | | |
|-------------------------|---------------------------------|
| BEL - Belgium | AFR - Africa |
| DEN - Denmark | ASI - Asia, Japan excluded |
| FRA - France | AUS - Australia and Pacific |
| FRG - Fed. Rep. Germany | CAN - Canada |
| ITA - Italy | CSA - Central and South America |
| NLD - Netherlands | EE - Eastern Europe |
| SWE - Sweden | JPN - Japan |
| SWI - Switzerland | RWE - Rest of Western Europe |
| UK - United Kingdom | USA - United States |

Source: *Tijssen and Moed*.²⁸

For the assessment of co-publication, addresses provide key information. The coverage of addresses varies strongly between databases: for 1984-1989, even after elimination of other items than articles, notes, letters and reviews, and those without references, about 57% of the items in the *A & HCI* carries an address, versus 90% in the *SSCI* and more than 95% in the *SCI*. Even if one is willing to assume that most

international co-publications appear in English, one has to face the fact that less than 68% of the English language articles carry an address in the *A & HCI*. Thus, especially in the *A & HCI*, trans-national publication rates are likely to be underestimated by the bad coverage of addresses.

The limitation to journal publications limits the usefulness of particularly the *SSCI* and *A & HCI*, as publications of books and chapters are important in the social and behavioural sciences and the humanities.^{29,30} However, the three ISI databases cover book publications only occasionally (mostly annual reviews). Therefore, they give only a limited indication of total trans-national productivity in the social and behavioural sciences and the humanities. In the *SSCI* and the *A & HCI*, the "internationality" of the publications is over-emphasized, as many fields are strongly locally oriented, causing scientists and scholars to publish a significant part of their work in local journals in local languages.^{30,31} Especially in non-Anglo-Saxon countries, these local journals are often not covered by *SSCI* or *A & HCI*. Thus, a significant part of the national output may not be visible in the *A & HCI* and the *SSCI*. These limitations are probably less severe in *SCI* studies.²¹

4. Conclusions and perspectives for future research

In this paper, we discussed several aspects of international scientific co-operation and awareness, two main elements of integration among scientists from different countries. This phenomenon is increasingly important, both for science policy makers and for scientists themselves. It is no wonder that it is the subject of various studies. Some of these studies consist of quantitative analyses, others are mainly theoretical or merely qualitative. At CWTS, several bibliometric studies have been performed. For panels set up by the CEC to evaluate its programmes on agricultural research and on medical and health research, we carried out bibliometric analyses in order to establish whether scientific co-operation and awareness within the EC has increased. The outcomes were that there has been an increase of intra-EC collaborations, but this growth is hardly faster than that of trans-national publications in general. For awareness as measured by citations, the same conclusions hold as far as agricultural research is concerned.

To infer that the EC programmes have failed in bringing about a considerable increase of international scientific co-operation and awareness would be premature. The phenomenon under study is very complex. All kinds of factors affect ISC. For instance, there are various forms of international co-operation, not all leading to

trans-national publications and before such publications become visible in international literature databases such as the *SCI*, a considerable amount of time has elapsed since the collaboration started. The patterns of international scientific co-operation and awareness seem to differ per country and per scientific (sub)field. In fields like astronomy ISC is strong, against very low rates of trans-national publications for most disciplines in the humanities. More insight is needed into the motives of scientists to co-operate internationally: complementary expertise and equipment, cost-sharing, the existence of stimulating structures and a favourable scientific climate, a high degree of mobility, and so on. With such an insight, models may be derived from behaviour of scientists in fields with a high rate of trans-national publications. Moreover, the validity of international co-authorships as indicator of ISC may be examined more closely.

The problem of validity of operationalization of theoretical concepts can be summarized in the following question: Is there equivalence in meaning between the concept-as-intended (international scientific co-operation) and the empirical variable-as-determined (international co-authorships)? One of the approaches to this validity problem is to examine the phenomenon of ISC more closely, distinguishing between various forms of scientific co-operation and analyzing whether such forms have resulted in international co-authorships. Another approach would be to 'start' from international co-authorships and to examine the relationship between the factors leading to such co-authorships and ISC.

The differences between countries with respect to international scientific integration are affected by both the policies of the national governments and long-term traditions in the political, economic, and cultural field. The decrease of British participation in some fields may be the consequence of national policy. Good examples of long-term factors are the strong ties of the United Kingdom and France with their former colonies, the co-operation networks among the Scandinavian countries, and the orientation of Eastern European countries on the Warsaw Pact and its allies in general and on the USSR in particular. Furthering ISC by the CEC is only one of the factors affecting the development of scientific integration among nations. Insight into the whole pattern of integration is extremely important for science policy makers (both on EC and on national level). To provide such an insight, more research is needed on the various factors affecting international scientific integration.

Other important issues relate to the effects of international scientific co-operation. Does ISC lead to a higher qualitative level of national research

performance? *Lewison and Cunningham*⁸ pose a different though related question: do trans-national papers have a higher impact than other papers?

An in-depth quantitative assessment of a phenomenon as complex as ISC may require sophisticated data analytical techniques. 'Maps' can be very useful in presenting findings of such techniques. These spatial configurations constitute a pattern recognition technique that identifies and depicts significant features of ISC structures in an understandable manner. Not only can ISC maps provide a better understanding of the underlying structure of international scientific cooperation, such maps also offer a different view of data which may yield new insights. This, in turn, may generate new focal points of interest in science policy issues concerning ISC. The concept of 'mapping' ISC relations can be extended to any array of quantitative data, and any combination of row and column elements (nations, scientific fields, institutes, researchers, etc.).

Tabular ISC data and ISC maps are primarily meant to reveal ('explore') the data; no statistical valid conclusions can be drawn. Although the data analytical methods that yield the ISC maps are accompanied by certain measures of inference, specific questions about the data may require a more thorough statistical analysis. One may think of issues such as: are there statistically significant differences in the ISC relations between a certain pairs of countries, or is there a significant trend in the ISC increase for some (pairs of) countries, or in certain scientific areas? Answers to such questions require an investigation into (or, perhaps, development of) statistical models of the stochastic processes that underlay ISC behaviour. So far, this topic has received less attention than it deserves.

Finally, an interesting problem is to examine the implications of the steady increase of the level of scientific co-operation between research groups (either at a national or at an international level) for the measurement of research performance of groups or countries. If a group publishes a high share (e.g., 50%) of its papers together with other groups, how can one assess the performance of such a group isolated from the performance of the co-publishing groups? A similar question may be posed for the measurement of research performance at a national level. Ranking of individual groups or countries based on publication and citation counts seems to become less and less appropriate tools for such a task, regardless of whether 'integer', 'fractional' or 'first-author' counting is applied.

In this paper, we emphasized the importance of our research for science policy, particularly in the EC frame-work. However, also for scientists on the working floor insight into patterns of international scientific co-operation and awareness is useful.

Being aware of the common interests of scientists within the European Community or having a clear picture of what exactly is going on in Eastern Europe, may be highly stimulating on scientific research. And last but not least, the questions posed in this paper are highly interesting in the field of science studies.

References

1. M.B. WALLERSTEIN (Ed.) *Scientific and Technological Cooperation among Industrialized Countries. The Role of the United States*, National Academy Press, Washington D.C., 1984.
2. G. LEWISON, The Evaluation of European Community Agricultural Research. Paper presented at the Conference of the Society for the Social Studies of Science, Irvine, USA, 1989.
3. M. O'SULLIVAN, A. FAVE, E. GALANTE, P. HENNING PETERSEN, W. HENRICHSMEYER, E. SANCHEZ MONGE, G. WILSON, Evaluation of the Agricultural Research Programmes (1976-1978, 1979-1983, and 1984-1988). Research Evaluation Report no. 39. Office for Official Publications of the European Communities. Brussels, 1989.
4. H.F. MOED, R.E. DE BRUIN, A. STRAATHOF, A Bibliometric Study of Agricultural Research within the European Community. Research Report for the Commission of the European Communities, Report 89-03, Centre for Science and Technology Studies, Leiden, 1989.
5. H.F. MOED, R.E. DE BRUIN, A. STRAATHOF, A Bibliometric Study for the Evaluation of the European Communities' Programmes on Medical and Health Research. Research Report for the Commission of the European Communities, Report 90-02, Centre for Science and Technology Studies, Leiden, 1990.
6. J.D. FRAME, M.P. CARPENTER, International research collaboration, *Social Studies of Science*, 9 (1979) 481-497.
7. L.A. LOMNITZ, M.W. REES, L. CAMEO, Publication and referencing patterns in a Mexican research institute, *Social Studies of Science*, 17 (1987) 115-133.
8. G. LEWISON, P. CUNNINGHAM, (1989), The use of bibliometrics in the evaluation of community biotechnology research programmes. In: A.F.J. VAN RAAN, A.J. NEDERHOF, H.F. MOED (Eds) *Science and Technology Indicators. Their Use in Science Policy and their Role in Science Studies*, DSWO Press, Leiden, 1989, pp. 99-114.
9. T. SCHOTT, International influence in science: Beyond center and periphery, *Social Science Research*, 17 (1988) 219-238.
10. M.H. MACROBERTS, B.R. MACROBERTS, Qualitative measures of communications in science: A study of the formal level, *Social Studies of Science*, 16 (1986) 151-172.
11. H.F. MOED, R.E. DE BRUIN, International scientific cooperation and awareness: A bibliometric case study on agricultural research within the European Community. *Newbury Park*, 1990, pp. 217-234, In: C.L. BORGMAN (Ed.), *Scholarly Communications and Bibliometrics*, Sage, London.
12. E. GARFIELD, *Citation Indexing. Its Theory and Applications in Science, Technology and Humanities*, John Wiley and Sons, New York, 1979.
13. R.E. DE BRUIN, H.F. MOED, The unification of addresses in scientific publications. In: L. EGGHE, R. ROUSSEAU (Eds), *Informetrics* Vol. 2, Elsevier Science Publishers, Amsterdam, 1990, pp. 65-78.
14. R. ARVANTIS, Y. CHATELIN, National strategies in tropical soil science. *Social Studies of Science*, 18 (1988) 113-46.
15. M.J. MORAVCSIK, The Coverage of science in the Third World: The "Philadelphia Program". In: L. EGGHE, R. ROUSSEAU (Eds), *Informetrics 87/88*, Elsevier Science Publishers, Amsterdam, 1988, pp. 147-155.

16. M.P. CARPENTER, F. NARIN, The adequacy of the Science Citation Index (SCI) as indicator of international research activity. *Journal of the American Society for Information Science*, (November 1981) 430-439.
17. R.A. ZWAAN, R.E. DE BRUIN, H.F. MOED, Publication Activity and Internationality of Dutch Research in the Social Sciences: A Methodological Discussion and Some Preliminary Results of an Online Bibliometric Study. Research Report for the Netherlands Advisory Council for Science Policy Research (RAWB), Report SSU-88-04, Centre for Science and Technology Studies, Leiden, 1988.
18. H.F. MOED, The use of online databases for bibliometric analysis. In: L. EGGHE, R. ROUSSEAU (Eds) *Informetrics 87/88*, Elsevier Science Publishers, Amsterdam, 1988, pp. 133-146.
19. H.F. MOED, New developments in online bibliometric analysis, In: A.F.J. VAN RAAN, A.J. NEDERHOF, H.F. MOED (Eds), *Science and Technology Indicators: Their Use in Science policy and their Role in Science Studies*, DSWO-Press, Leiden, 1989, pp. 115-127.
20. H.F. MOED, Publication Activity and Internationality of Dutch Research in the Natural and Life Sciences (1984-1987). Report SSU-88-03, Centre for Science and Technology Studies, Leiden, 1988.
21. A.J. NEDERHOF, H.F. MOED, A.M.P.H. RAMAEKERS, Trends in Publication Activity and International Co-publication in the Natural and Life Sciences (1984-1989): Centre for Studies of Science and Technology, Leiden, 1990.
22. A.J. NEDERHOF, E.C.M. NOYONS, Trends in Publication Activity and International Co-publication in the Social and Behavioral Sciences, and the Humanities (1984-1989), Centre for Studies of Science and Technology, Leiden, 1990.
23. R.E. DE BRUIN, H.F. MOED, International Scientific Co-operation in European Scientific Research. An Historical Interpretation. Paper Presented at the Conference of the Society for the Social Studies of Science, 1990.
24. M. STANILAND, Francophone Africa: The enduring French connection, *Annals of the American Academy of Political and Social Science*, 489 (1987) 51-62.
25. D. MOOSI, French foreign policy: The challenge of adaptation, *Foreign Affairs*, 67/1 (1988) 151-164.
26. L.R. LUBBRANO, National and international politics in US-USSR scientific cooperation, *Social Studies of Science*, 11 (1981) 451-480.
27. M. GREENACRE, *Theory and Applications of Correspondence Analysis*, Academic Press, New York, 1984.
28. R.J.W. TIJSSEN, H.F. MOED, Literature-based statistical analyses of international scientific co-operation: An exploratory case study of the Netherlands. In: A.F.J. VAN RAAN, A.J. NEDERHOF, H.F. MOED (Eds), *Science and Technology Indicators: Their Use in Science policy and their Role in Science Studies*, DSWO-Press, Leiden 1989, pp. 129-146.
29. A.J. NEDERHOF, Books and chapters are not to be neglected in measuring research productivity, *American Psychologist*, 44 (1989) 985-986.
30. E.C.M. NOYONS, A.J. NEDERHOF, International Comparisons of Departments Research Performance and Assessment of External Knowledge Transfer in the Humanities and the Social and Behavioral Sciences. Report CWTS-89-04, Centre for Science and Technology Studies, Leiden, 1989.
31. A.J. NEDERHOF, R.A. ZWAAN, Quality judgements of journals as indicators of research performance in the Humanities and the Social and Behavioral Sciences, *Journal of the American Society for Information Science*, in press.