

Studying the brain drain: Can bibliometric methods help?

GRIT LAUDEL

*Research Evaluation and Policy Project, Research School of Social Sciences, Australian National University,
Canberra (Australia)*

Today science policy makers in many countries worry about a brain drain, i.e., about permanently losing their best scientists to other countries. However, such a brain drain has proven to be difficult to measure. This article reports a test of bibliometric methods that could possibly be used to study the brain drain on the micro-level. An investigation of elite mobility must solve the three methodological problems of delineating a specialty, identifying a specialty's elite and identifying international mobility and migration. The first two problems were preliminarily solved by combining participant lists from elite conferences (Gordon conferences) and citation data. Mobility was measured by using the address information of publication databases. The delineation of specialties has been identified as the crucial problem in studying elite mobility on the micro-level. Policy concerns of a brain drain were confirmed by measuring the mobility of the biomedical Angiotensin specialty.

Introduction

The interorganisational mobility of scientists has always been an important functional requirement for science. Scientists “on the move” bring their knowledge to other places, acquire new knowledge in the new place and thus promote new combinations of knowledge. This is especially important if knowledge is not communicated through other channels like publications (e.g., if tacit knowledge is included). Since some kinds of knowledge are circulated in science mainly by scientists who travel around, scientists' interorganisational mobility constitutes one of the most important knowledge flows in science.

Spatial mobility is a widespread phenomenon in science. In the course of their career, scientists usually work in several organisations. For example, many future scientists acquire their PhD at a different university to which they studied, and scientists usually leave the university at which they received their PhD. In later career stages, Fellowships allow temporary stays at other organisations. All these forms of organisational mobility are usually open to scientists from all countries. International advertisements of positions and internationally composed research groups have been

Received March 27, 2003

Address for correspondence:

GRIT LAUDEL

Research Evaluation and Policy Project, Research School of Social Sciences

The Australian National University, Canberra, ACT0200, Australia

E-mail: grit.laudel@anu.edu.au

0138–9130/2003/US \$ 20.00

Copyright © 2003 Akadémiai Kiadó, Budapest

All rights reserved

common in science long before globalisation entered into economic practices and discourses. In addition, it is meanwhile a standard practise in some sciences that young scientists from Europe have a postdoctoral career phase in the USA and then return to their home countries.

Sometimes, scientists move to other countries and stay there for the rest of their career. In this case, mobility becomes migration, i.e., a scientist's permanent move to a country different from his or her country of origin. This phenomenon causes migration currents between countries, which can be described by balances of "gaining" and "losing" scientists.

Negative balances are always a point of concern because the education of scientists is expensive. However, there is a specific quality of losses that is troubling current science policy: A national science system can be said to be in trouble when it loses its best scientists to other countries. We then talk about a brain drain, that is the migration of elite scientists out of one country. From the point of view of the country to which the elite scientists migrate it is a brain gain. Recently, many countries are concerned about losing their elite scientists to other countries, especially to the USA (e.g., SINCELL, 2000; HELLEMANS, 2001; BMBF, 2001: 6-7). This fear has already led to several policy measures of "buying back" elite scientists with extraordinary grants.* However, the concerns about brain drains are mainly based on anecdotal evidence. There is little data about the extent of brain drains or about affected fields. Two recent studies that were initiated by German science policy to study the brain drain of German academics could not solve the problem either (BUECHTEMANN, 2001; STIFTERVERBAND FÜR DIE DEUTSCHE WISSENSCHAFT, 2002). In the 2001 study only cautious estimations on the basis of plausibility arguments are made that there seems to be a brain drain, especially in new and interdisciplinary research fields. The authors of the 2002 study summarize their results as follows: "It is questionable if the term Brain Drain suitably describes the recent German situation. At present, valid statements about the extent of a permanent or temporary migration of German scientists abroad or the movement of foreign scientists and highly qualified specialists into Germany cannot be made" (STIFTERVERBAND FÜR DIE DEUTSCHE WISSENSCHAFT, 2002: 1, my translation). They see one of the difficulties as the lack of appropriate data on population statistics (ibid: 8-9). But the problem cannot be solved by just improving population statistics because it is rooted in the

* In 2001, the Australian Research Council created a new funding scheme, the "Federation Fellowships". The aim was to "provide opportunities for outstanding Australian researchers to return to, or remain in, key positions in Australia" (AUSTRALIAN RESEARCH COUNCIL, 2001). Also in 2001, Germany's research minister announced extra money for a 'brain gain' of the best foreign scientists as well as German scientists working abroad (BMBF, 2001: 7). Similar policy measures were undertaken by other countries (e.g., PICKRELL, 2001; SPURGEON, 2000).

conceptual and methodological difficulties of measuring the mobility of elite scientists. The methods that have been used to measure mobility - questionnaires, or CV data from various sources – have serious shortcomings that hinder detailed analysis. An alternative approach that suggests itself because of the availability of the address information in bibliographic databases is to apply bibliometric methods. Surprisingly, these methods have not been systematically applied to the study of scientists' spatial mobility.

The aim of this paper is to test whether bibliometric methods can be used for the study of elite scientists' international mobility in general, and thus can be used for the study of the brain drain. In order to answer this question, three problems had to be solved:

- Delineating a specialty: Scientists are elite members in relation to the specialty to which they belong. These specialties have to be delineated.
- Identifying a specialty's elite: A specialty's elite are those scientists who make the most significant contributions to their specialty's knowledge production and provide an orientation for all specialty members.
- Identifying international mobility and migration: An elite researcher's movements from one country to another have to be analysed. It must be checked whether or not a movement is permanent.

Because the aim of this article is to initiate a methodical discussion, the description will also include some of the methodical failures encountered. For the methodical test a biomedical specialty was selected. The Results section presents the empirical results of this test study and discusses the methodological results of all the tests. Finally, methodological conclusions will be drawn about the applicability of bibliometric indicators for studying scientists' mobility.

State of the art: concepts and methods used for analysing scientists' mobility

Conceptual problems

The analysis of elite migration hinges on several important conceptual decisions. Firstly, an elite is always an elite of another, larger, group. This group needs clarification. The common notion of "scientific elite", which generally refers to a country's best scientists, is useful only for very general science policy analyses. In most cases, it should be of interest how a country's brain drain is distributed across scientific fields, and what causes the - probably uneven - distribution. Causes of elite migration

can be assumed to be complex and partly related to conditions of research as well as to the location of other elite members, e.g., to the existence of a “critical mass” of elite members in a scientist’s home country. This means that while elite migration appears to be a very general phenomenon, we cannot be sure that the country level does not obscure important differences between specialties. A variation of elite migration is likely to exist because the causes of this process are partly field-specific. The consequences of elite migration are also likely to vary between specialties. Sociology of science has repeatedly proposed that the scientific elite has a governing and regulating function. Losing a specialty’s elite members means that this national specialty becomes uncoupled from frontier science, that quality standards might no longer be enforced nationally, and that an important channel of communicating societal interests to those who govern the specialty gets lost. The often used but rather fuzzy concepts of “best scientists” or “scientifically excellent scientists” are not unequivocally related to a reference group. While a loss of an excellent scientist is clearly undesirable, its functional consequences for a national science system are difficult to describe.

While we all know that there is a brain drain towards the USA, we do not know if an elite migration occurs in all specialties, and we do not know to what extent the reason for elite migrations are field-specific. In order to identify the migration that can affect a country’s science functionally (elite migration) and in order to target field-specific causes for this migration, science policy needs to know the field specific causes and consequences of this process. Detailed investigations of brain drains can only be undertaken if the different scientific specialties’ elites are identified and investigated. That is why delineating specialties is an important task in analysing brain drains, a task that in turn requires a clarification of the concept “specialty”. A specialty is understood here as a community of scientists who directly or indirectly interact in the production of new knowledge about a common subject matter. This joint production is decentrally coordinated by the specialty’s members who interpret the specialty’s shared body of knowledge, derive problems and draw means for their solution from it, and propose how the solutions can be integrated in this body of knowledge (GLÄSER, 2001: 194-196). Specialties vary in their size and can be assumed to consist of any number between a dozen and several thousand.

Among the few studies that provide empirical data about scientists’ mobility, hardly any focus on the movement of the scientific elites. But all these studies had to cope with the problems of delineating research fields and how to measure mobility. Therefore, they are included in the discussion. I will now describe the methods that were used by others to measure scientists’ mobility and discuss their shortcomings, following the three steps described above.

Delineating a specialty

Most of the mobility studies differentiate only very roughly between groups of research fields. STEPHAN & LEVIN (2001: 61-63) distinguish earth/environmental sciences, life sciences, mathematical and computer sciences, physical sciences, and engineering. MAHROUM (1999: 383) differentiates between clinical medicine, biosciences, chemistry, mathematics, physics etc. In other mobility studies, similarly broad fields are delineated (e.g., VAN HEERINGEN & DIJKWEL 1987; SHAUMAN & YU, 1996). Others did not delineate between different research fields at all (e.g., PIERSON & COTGREAVE, 2000).

TRIMBLE (2000) studied the brain gain of a sample of American astronomers and astrophysicists by tracking their country of origin. She selected a sample of young astronomers (applicants for tenure-track positions at two American astronomy departments and applicants for the American part of the International Astronomical Union) as well as established astronomers (officers, councillors, and committee members of the American Astronomical Society and a subsample of the International Astronomical Union). This selection covers only a part of American astronomy. Furthermore, astronomy can be assumed to consist of several specialties, an internal structure that was neglected by Trimble. ROSENFELD & JONES (1987) investigated the mobility of American psychologists using data of the American Psychological Association. The study has two delineation problems: firstly, it did not distinguish between members of the association who are scientists and members who are practitioners; secondly, the divisions of the association indicate that it includes more than one psychological specialty, e.g., history and philosophy of psychology, clinical psychology, and neuroscience.

The studies on mobility that have been conducted so far did not pay attention to the problem of delineation. Often the authors started with data of all scientists of one country, and introduced the field delineation only as a second step, in nearly all cases ending up with entities that are much broader than a specialty. This is not a problem as long as only a description of mobility is intended. However, whenever causes and consequences of mobility, especially of elite mobility, are to be investigated, the delineation of a specialty is important. For example, delineation becomes a crucial step if working conditions (access to resources) and cumulative effects (elite members go where elite members are) are investigated in comparative studies between countries or specialties.

In contrast, bibliometricians have been preoccupied with the problem of the delineation of specialties for a long time. The easy way to achieve a delineation appears to be to apply the field classification (“subject category”) provided by the Institute of Scientific Information (ISI). More sophisticated methods use cluster analyses based on co-citations (e.g., SMALL, 1977) or filters for key words as well as words from titles, abstracts, or addresses (NOYONS & VAN RAAN, 1998; LEWISON, 1999; NOYONS, 2001) for mapping research fields. An important shortcoming of all these bibliometric methods is that they must apply arbitrary thresholds in their construction of boundaries between fields. Word-based approaches additionally include expert judgements (LEWISON, 1999; NOYONS et al., 2002), which may introduce the experts’ specific perspective on the field. Comparisons of co-citation analysis and co-word analysis show that the delineated research structures differ, as demonstrated by NOYONS et al. (2000: 157-209). It is difficult to specify what is delineated by co-citation analysis and what is delineated by co-word based methods. That is why citation-based and word-based methods, though current best practice, are still not completely satisfactory (VAN RAAN, 1997: 215; NOYONS & VAN RAAN, 1998: 77-80).

Identifying a specialty’s elite

Stephan & Levin provide the only study of the scientific mobility of elite scientists (LEVIN & STEPHAN, 1999; STEPHAN & LEVIN, 2001). They use the term “elite” related to a national science system rather than related to a specialty. However, even with this simplified concept they had the methodological problem of finding the “leading scientists” of all specialties situated in the USA. The authors used a number of methods to delineate the American scientific elite. Firstly, members of the US National Academy of Sciences and of the National Academy of Engineers were selected. As honorary members were excluded, the academy membership appeared to be a good indicator for a scientific elite if combined with other indicators. Academy members “are elected in recognition of their distinguished and continuing contributions to knowledge” (STEPHAN & LEVIN, 2001: 65). But since Academy membership is a lifetime award, among the Academy members are also scientists who belonged to the elite previously but are not active any more.*

* This problem could be solved by adding an analysis of each members’ recent publications and their citations. A short test shows that for example NAS Member Edward A. Adelberg had between 1990 and 2002 two publications in the ISI database (*Web of Science*), one in 1991 and one in 1998, and that the second publication is his scientific biography. So it cannot be assumed that he has still an elite function in his field.

The other indicators applied by Stephan and Levin are bibliometric indicators. Three of them are discussed because they are potentially useful for solving our problem. The first indicator is the so-called “citation classics”. That is highly cited journal articles indexed in the ISI databases that attain a certain citation threshold. Secondly, “hot papers” were included. Hot papers are “journal articles published during the most recent two-year period that in the most recent two-month period have attracted significantly more attention than papers of the same age in the same field” (<http://www.isinet.com/isi/products/rsg/products/sw-hp/>).

Since in the sciences most of the papers are co-authored, neither indicator necessarily measures the quality of a single scientist. The indicator always refers to the small group of coauthors. Stephan and Levin distinguish between the first author and non-first authors of a “citation classic” or a “hot paper” and present data for both types of authors. Obviously, the authors attempted to differentiate between important and less important contributions of the coauthors that usually exist. However, the chosen approach is questionable because it presupposes the existence of highly specific rules about name ordering. Two objections can be raised. Firstly it is by no means sure that the shift from alphabetical to contribution-based name ordering has taken place in all specialties. No comprehensive empirical investigation across all fields has been conducted. Secondly, even in the case of a contribution-based name ordering significant contributions may stem from scientists other than first authors. In a study on the rewarding of collaborative contributions I have shown how contributions of different types are rewarded in several fields (LAUDEL, 2001). Apparently there is no justification for disregarding all scientists except the first author. Owing to the division of labour within the sciences, the first author usually has done the experimental work and the last author the theoretical-conceptual work. If theoretical contributions from two collaborating groups are combined, then the two last authors can be expected to be the group leaders who made these contributions. These general rules are modified by other rules and factors (ibid). Generally speaking, the name-ordering of coauthors is not reliable enough to enable unambiguous conclusions about the importance of the coauthors’ contributions. For these reasons it is impossible to draw conclusions from multi-authored papers about the elite status of one author. Moreover, Hot Papers only indicate a single outstanding contribution. This is not sufficient for classifying a scientist as an elite member because an elite scientist who is providing an orientation to other scientists of his or her specialty can be assumed to do this with more than one significant contribution.

A third bibliometric indicator applied by Stephan and Levin is “most-cited scientists” during the years 1981 to 1990 as provided by ISI (Essential Science Indicators). This is also a problematic strategy because one of the major shortcomings of these ISI data is that homonyms are not corrected.* The errors described above might statistically vanish in the study of Stephan and Levin because it is a study on the macro-level of science. But whenever elite mobility of single specialties or small fields is concerned, the weaknesses of SCI-based indicators cannot be ignored.

Söderqvist and Silverstein suggest using quantitative data of participation in scientific meetings for “identifying leaders of a scientific discipline and its subunits” (SÖDERQVIST & SILVERSTEIN, 1994: 243). Scientists who most frequently participate in meetings are identified as those leaders. The method fails because not all leading scientists in the investigated field (immunology) showed a high frequency in conference participation and not all scientists that frequently participate in conferences of their field could be assigned as leading scientists (ibid: 247).

Elite members were also identified by using sociometric techniques, based on questionnaires. Crawford asked sleep researchers “to name all persons they contacted at least three times during the past year concerning their work” (CRAWFORD, 1971: 303). Amick asked chemists from a special geographical region to identify the person who is, in their opinion, the outstanding chemist in this region (AMICK, 1974: 3). These techniques, like co-citation and co-word analyses, require the introduction of arbitrary thresholds to delineate elite members from non-elite members.

With the growing number of evaluations of research performances, methodological questions arose about the way bibliometric indicators measure scientific excellence (VAN RAAN, 2000; TIJSEN et al., 2002). The term “scientific excellence” seems to be so common that it is not further defined in the named articles. On the level of individual researchers it can be equated to being an elite scientist, that is: making the most significant contributions to his or her specialty and providing an orientation for all specialty members. Scientific excellence on the individual level can be analysed by using elaborated citation measurements (VAN RAAN, 2000: 305-309). A limitation of the proposed indicators is that their statistical significance may be endangered when the

* “Authors having the same last name and initials may represent multiple individuals. This is especially likely in the case of common surnames. The ability to breakout the name by field may to some degree disambiguate person X in field Y from person X in field Z, however, keep in mind that a listed name can still represent more than one author within the same field.” Furthermore a scientists’ name can be written in different ways – with and without middle initials for example. The scientist then appears as two persons (ISI Essential Science Indicators Online help, Scientists, Scientists Data Information – <http://essentialscience.com>). Stephan and Levin were aware of the homonym problem and worked partly around it by eliminating very common names (STEPHAN & LEVIN, 2001: 66 and footnote 12).

number of a scientist's publications is low (ibid: 309). Citations can indicate that a scientist made substantial contributions in his or her field. What cannot be analysed by mere citation counts is the orienting function of the elite member for other specialty members. Therefore, citation-based indicators must be carefully designed if they are used to delineate a specialty's elite.

Identifying international mobility and migration

In order to measure the interorganisational mobility of scientists several techniques have been used. Most of the techniques were aimed at getting information about the curriculum vitae of the scientists.

- *Encyclopaedias of scientists*: Encyclopaedias such as "American men and women in science" were used (ALLISON & LONG, 1987; TRIMBLE, 2000; STEPHAN & LEVIN, 2001). They contain CV data for the scientists included.
- *Biographical information provided by Scientific Associations* were used, e.g., by the American Psychological Association (ROSENFELD & JONES, 1987) or by the National Academy of Science (STEPHAN & LEVIN, 2001).
- Available *data of population statistics* were applied, e.g., American census data and surveys (SHAUMAN & YU, 1996; BUECHTEMANN, 2001), Statistical data of the Nordic countries (GRAVERSEN et al., 2002), Statistical data of the British Higher Education Statistics Agency (MAHROUM, 1999).
- Curriculum vitae (CVs) that are an integral part of grant applications were used. TRIMBLE (2000) used the CVs of applicants for tenure track positions as well as for the membership in the International Astronomical Union. MARTIN et al. (1996) used information in grant applications of a Spanish Fellowship Programme. In the study of the STIFTERVERBAND FÜR DIE DEUTSCHE WISSENSCHAFT (2002), data from funding agencies with international exchange programmes were used that keep addresses of the funded research fellows.
- *Questionnaires* were sent out by mail or email to obtain CV data (STEPHAN & LEVIN, 2001; VAN HEERINGEN & DIJKWEL, 1987; DEBACKERE & RAPPA, 1995).
- *CV data available from the Internet*: DIETZ et al. (2000) did a general methodical test for getting CV data from the Internet.

These techniques might have been useful in the context in which they were applied. But in relation to the aim of measuring mobility of a specialty's elite scientists, they have three important shortcomings. Firstly, they provide incomplete data. A study on

elite mobility that is conducted at the level of a specialty strongly requires complete data, i.e. mobility data about all elite members identified. The number of elite scientists within one specialty is very small. Therefore, even a small amount of missing data will have strong effects. None of the abovementioned techniques can guarantee completeness. The problem is well-known for questionnaires: The response rate in studies of mobility varied between 38.4% and 54.5%. Response rates are partly affected by the problem under study: Mailed questionnaires about mobility are aimed at a moving target. A new opportunity to study mobility is provided by the internet. However, internet data are not complete at all because it is usually the scientist who decides about the presentation of their CV data on the websites. Some scientists simply do not publish their CVs, while others provide only limited data. Encyclopaedias are unlikely to contain data about all scientists who have been identified as elite members.

Other data do not depend on scientists' behaviour but still may suffer from incompleteness. For example, biographical data may be provided only for scientists from a certain geographical area, as is the case with encyclopaedias, CVs from members of scientific associations, or population statistics. Since specialties are entities that cross the boundaries of one country this is a serious limitation for mobility studies of specialties.

A second obstacle to studying elite mobility is the limited access to CV data provided in grant applications or in membership lists of scientific associations. Neither funding agencies nor scientific associations easily grant access to CV data. Since usually more than one of these organizations must be approached if complete data about an elite's mobility are to be obtained, a study may be seriously hindered by incomplete access.

Both obstacles to obtaining sufficient data for a study of mobility are caused by the fact that there is a third party who governs the access to the data. This difficulty can be circumvented by using bibliometric methods. PIERSON & COTGREAVE (2000) did use SCI data to find out the stay of researchers who obtained a doctorate in a science subject from a British university and are still active in science, using the ISI address field. In their very short description they don't give many methodological details how they solved certain measurement problems (e.g., with homonyms or erroneous addresses). Beyond that, no methodological discussion of the application of bibliometric methods for mobility measurements has taken place so far. Therefore I decided to test them systematically.

Methodical tests of measuring elite scientists' mobility

The test will follow the three steps of delineating a specialty, identifying the elite, and measuring the mobility. For a better judgement of the proposed methods, I will also include the failures. Moreover, they make it more understandable why I chose a rather unusual way.

Delineating a specialty

Using ISI field classification. I started with one of the common methods in bibliometrics by using the field classification provided by ISI. Within its instrument "Essential Science Indicators" ISI provides 22 research fields, such as physics, chemistry, materials science. ISI has assigned the journals it covers to so-called subject categories, 170 in the Sciences and 54 in the Social Sciences (own calculations from the *Journal Citation Report*). Both research field classifications are based on journal-to-journal citations, which are frequently used for delineation purposes in scientometrics. Since a journal usually contains articles from more than one specialty, ISI's classifications are too broad to depict entities as small as scientific specialties. This limitation of the subject categories has been observed in an earlier investigation of an interdisciplinary collaboration network. Most of the research groups' publications were assigned to the categories Biochemistry & Molecular Biology or Biophysics. But in the interviews cognitive distances as well as communication problems were mentioned that were disguised by the scientists' apparent belonging to the same field. The subject classification system proved to be too rough to delineate a specialty (LAUDEL, 1999: 66). Furthermore, the field delineation is too inflexible to depict the dynamics of scientific specialties: There is a constant process of the emergence of new specialties by integration and disintegration processes that ISI cannot keep up with its subject categories. For these reasons, ISI field classification is not suitable to delineate scientific specialties (see also AKSNES et al., 2000).

Using conference participant data. In my interviews with scientists in previous projects I learned that scientists sometimes attend highly specialised conferences that are aimed at bringing together members of one specialty. These conferences seem to depict the current status of a specialty much better than the bibliometric field classification discussed above. In order to solve the problems of delineating a specialty and identifying its elite simultaneously, I chose elite conferences, namely the Gordon conferences, as my starting point.

Gordon conferences are highly prestigious and specialised conferences. The number of carefully selected participants is limited to 135. Scientists who have made one or several outstanding contributions to the field are invited to give lectures. Besides this speaker role, there are several other roles at the conference: chair, vice chair, discussion leader, poster presenter and attendee. All these forms of participation presuppose that the scientists have been chosen by a selection committee. The lists of participants are available on the websites of the Gordon Conference Organisers (<http://www.grc.uri.edu>). A participant list consists of the scientists' name, their institutional affiliation and the form of participation.

Gordon conferences vary with regard to the scope of their subject. Some themes (e.g., Analytical Chemistry) appear to be too broad to encompass only one specialty. I selected the "Angiotensin" Gordon Conferences. "Angiotensin" conferences present a relatively easily identifiable biomedical specialty, namely the specialty that is formed around a substance (Angiotensin) that is the specialty's common research object. Scientists of this specialty were selected from the six available participant lists within the time span of 1996 to 2002 (there was no conference in 2000).

ANGIOTENSIN			
Ventura Beach Marriott Mar 11-16, 2001			
This was the final registration list for the 2001 Conference on ANGIOTENSIN.			
Name	Organization	Participation	Status
LEN ADAM	BRISTOL MYERS SQUIBB	Attendee	Registered
WAYNE ALEXANDER	EMORY UNIVERSITY HOSPITAL	Discussion Leader	Registered
MOHAMMED SHOWKAT ALI	EMORY UNIVERSITY	Attendee	Registered
MARIA CLAUDINA ANDRADE	UNIVERSIDADE FEDERAL DE SAO PAULO	Attendee	Registered
KENNETH M BAKER	TEXAS A&M HSC	Discussion Leader	Registered
ANTHONY J BALMFORTH	UNIVERSITY OF LEEDS	Attendee	Registered
MATTHIAS BARTON	UNIVERSITY HOSPITAL, ZURICH	Poster Presenter	Registered
BRADFORD C BERK	UNIVERSITY OF ROCHESTER MEDICAL CENTER	Chair	Registered

Figure 1. Participant list of the Angiotensin Gordon conference in 2001 (extract)

The Gordon conferences have a strong American tradition. Three of the Angiotensin Gordon Conferences in the investigated time-span took place in the USA, the three others in Europe. The fact that half of the conferences were in the United States might

indicate a US bias. This was checked by testing the following assumption: If there is an American bias, American participants are invited more often to Gordon conferences despite a lower scientific performance, which means in turn that the average performance of non-American participants is higher. I checked for the 215 scientists from the reduced participant list (Figure 1) whether they had either a “hot paper” or belong to the Top 1% scientists in the ISI Essential Science Indicators. The Chi-Square-test showed that there are no significant differences between American participants and non-American participants.

Identifying a specialty's elite

Using ISI's “research fronts”. An attempt to identify an elite independently from the Gordon conferences was the application of ISI's “research fronts”. ISI provides in its instrument Essential Science Indicators co-citation clusters of highly cited papers, so-called “research fronts”. “A research front is a group of highly cited papers, referred to as ‘core papers’, in a specialized topic defined by a cluster analysis. ... A measure of association between highly cited papers is used to form the clusters. ... That measure is the number of times pairs of papers have been co-cited, that is, the number of later papers that have cited both of them. Clusters are formed by selecting all papers that can be linked together by a specified co-citation threshold” (http://www.isinet.com/demos/esi/h_datres.htm).

This instrument apparently solves the problem of field delineation because research fronts are much smaller than the “fields” provided by ISI. However, what resulted was that co-citation clusters recall only a certain proportion of a specialty's output and therefore cannot identify all of its members. Moreover, the elements of a research front – highly co-cited journal articles – suffer from the problems described earlier. Most of the articles have more than one author, not all of which can be assumed to belong to a specialty's elite.

Reducing the list of Gordon conference participants. The list of participants in Gordon conferences provided already “more than average” scientists of the field. However, it was doubtful that the specialty has so numerous an elite as represented by the list of all scientists who ever attended a conference. Therefore, the list had to be reduced, and the “true elite” had to be identified among the scientists on the list.

Table 1. Reduced list of 215 active participants in the Gordon conferences (extract)

Name	Organization	Participation
AGNES FOGO	VANDERBILT UNIVERSITY	Speaker
AKIYOSHI FUKAMIZU	UNIVERSITY OF TSUKUBA	Speaker
ALAN DAUGHERTY	UNIVERSITY OF KENTUCKY	Speaker
ALBERTO NASJLETTI	MEDICAL COLLEGE OF NEW YORK	Speaker
ALESSANDRO M CAPPONI	UNIVERSITY HOSPITAL	Chair
ALISTAIR V FERGUSON	QUEEN'S UNIVERSITY	Speaker
ANDREW S GREENE	MEDICAL COLLEGE OF WISCONSIN	Speaker

In a first step, active members above the poster threshold were selected. Chairs, Vice Chairs, Discussion leaders and Speakers were included as active participants. Poster presenters were omitted because posters are usually the less important contributions. Although the attendees were also selected by the Gordon Conferences' scientific committee, I excluded them for having the weakest contributions. The remaining list contained 215 active participants. In order to further reduce this list, I applied several bibliometric techniques.

Using ISI Top 1% classification of scientists. The participant list contained scientists of the field who had made one or more outstanding contributions. An elite scientist was defined as a person who is continuously orienting the work of his or her specialty, i.e., as making more than one outstanding contribution. Therefore, the scientists with only one outstanding contribution had to be eliminated. The idea was to find the most highly cited scientists.

I used the ISI Essential Science Indicators data of the top 1% of scientists. This list rests on citation counts within the 22 broadly defined research fields. Each of the fields has a certain citation threshold. The use of this data failed for two reasons: Firstly, the Angiotensin scientists were found in four of the 22 fields (Biology & Biochemistry, Clinical Medicine, Molecular Biology & Genetics, Multidisciplinary). These four fields have different citation thresholds for including or excluding scientists. If a scientist is assigned to the "wrong" ISI field s/he might fail the inclusion in the top 1%. Sometimes a scientist works on the border of several ISI fields and is highly cited in all of them. If s/he is unlucky the sum of all citations could be high but the citations in one of the 22 ISI fields was not high enough to be included in the top 1% of highly cited scientists. Secondly, ISI does not control in these indicators for homonyms and does not provide sufficient information for eliminating them.

Using ISI citation scores. The easy way to use the calculations that have already been done by ISI had failed. Therefore I decided to calculate the citation scores of the 215 selected Angiotensin researchers myself. I counted the citations per publication of

each scientist for a 10 year time-span. A major shortcoming of this method is that citations are counted irrespective of the citing specialty. Scientists publications are cited not only within their specialty but also in “neighbouring” specialties. Therefore, general citation counts can be misleading. In order to identify an elite role in relation to a specialty, only citations from specialty members should be counted. This was, however, impossible because the participant lists of Gordon conferences were not sufficient for delineating the whole specialty.

Analysing citations and cocitations within the conference participants sample. Ideally, one would count only citations coming from a scientist’s own specialty. This was impossible due to the delineation problems described in the previous section. As a substitute, the sample of 215 Angiotensin researchers who actively participated in Gordon conferences was used as a reference group in order to calculate “internal” citation and co-citation scores.

All publications of the 215 scientists between 1990 and 2002 were downloaded from ISI’s *Web of Science*. Homonyms were eliminated by analysing address fields, journal titles, titles of articles and searching the internet. Meeting abstracts, notes and letters were eliminated. The remaining 13020 articles were analysed using Microsoft VBA Macros. Further cleaning of data was necessary to standardise journal titles (titles used in publications differed from titles used in citations) and authors’ names (in some cases, hyphens had to be deleted from author names because they occur in author names after 1997 but not in citations).

For all 215 authors, citations by the other 214 authors of the samples were counted, and the authors were ranked according to their citations by the sample. Thereafter, an author cocitation analysis within the sample was conducted (no thresholds for citations or cocitations were used). Adopting an idea of NOYONS & VAN RAAN (1995; 1998: 9), the relation between thresholds and cocitation cluster sizes was investigated. Since no plateau or curvature was discovered that would justify the introduction of a threshold for delineating the specialty’s elite, authors were ranked according to the number of authors they were cocited with. This indicator stems from network analysis, where the degree of a node is calculated as the number of links with other nodes in the network (SCOTT, 1991: 70).

The first 150 places of the internal citation rank list and of the author cocitation rank list were compared, and authors who ranked 150 and better on both lists were selected as the likely elite. For these 131 scientists, a test of mobility measures was conducted. The selection of the scientists with the highest scores on both the “internal” citation list and the “internal” author cocitation list was an arbitrary decision. As has been indicated above, a detailed analysis of orienting roles within the specialty Angiotensin was

impossible because of the general delineation problem. However, for purposes of a methodological test (and as a pilot study of the significance of the problem) this rough selection was sufficient.

In order to make sure that the procedure has led to the selection of better-than-average scientists, I compared the citation per publication scores of the 100 best Gordon conference participants with the citation per publication scores of some Australian Angiotensin researchers who were not participants of Gordon conferences.* The comparison showed that all the Australian non-participants had a clearly lower citation per publication score than the best 100 Gordon conference participants. Thus, the identification of a scientific elite rested on the combination of two independent methods. I started with the elite conferences and thus used the attribution of the specialty, i.e., its decisions about who has made outstanding contributions in its field. In a second step I identified the scientists with most co-citation links and with most citations within this group.

Identifying international mobility

Using address information in a journal publication database. The international mobility of the selected 131 Angiotensin scientists had to be identified. I assumed that a scientist's first publications are based on the PhD work – scientists who don't publish the results of their PhD thesis are very unlikely to pursue a career in science afterwards. Therefore, first publications occurring in databases were defined as indicating the starting point of both the career and the mobility history of a scientist.

I used the biomedical-oriented database PubMed to reconstruct the mobility histories of the 131 scientists. This database contains publications from the mid 1960's until now. It provides the institutional affiliation of the first author only since the mid 1980's. By selecting all first-authored publications of a scientist I could track the scientist's movements between institutions. If the scientist was not a first author on any publication over a couple of years, I looked up the address in the online version of the journal. Additionally, CV data was obtained from the internet in order to supplement bibliometric data. For scientists who published before the mid 80's and who had no CV in the Internet I looked up the address information in the hard copy version of the article in the journals. The internet CV data were also used to check if the changes in

* I used the Australian Publication database that consists of all ISI publications that have at least one Australian address in the address field. I found the scientists by using Angiotensin as title keyword. Only scientists who had at least five publications with this title keyword were included.

institutional affiliations given in the publications correspond to the true movements of a scientist. It showed that the bibliometric data reflected the movements of a scientist (usually with a time-delay of one year).

Comparing different publication databases for analysing interorganisational mobility. I tested two other publication databases to find out whether they enable mobility analyses. I compared *Pubmed* with the *Web of Science* that contains all research fields; as well as with *INSPEC* that contains articles in the field of physics and engineering. Table 2 shows the results of these tests.

Table 2. Comparison of three publication databases for supporting scientific mobility analyses

Feature	<i>Web of Science</i>	<i>PubMed</i>	<i>INSPEC</i>
Scope of fields	all	Biosciences	Physics and Engineering
Scope of years	Since 1990*	Since mid 1960s	Since 1969
Address information since	From the beginning	mid 1980s	1969
Institutional affiliation	Corresponding author's affiliation; all addresses but not directly linked to the author name	Only first authors' affiliation	Only first authors' affiliation
Identification of the authors	First name as initial	First name as initial	First name as initial
Homonyms	Frequent because of the coverage of all fields	Less frequent because of the limitation to biological fields	Less frequent because of the limitation to physical and engineering fields
Connection to the online full text version of the article	Sometimes in recent years	Sometimes in recent years	No
Download of the publication list	slow	fast	fast

*I used the access of the Australian National University that starts with 1990. The ISI database of the sciences goes back until 1945. The problem is that most institutions don't have the version that covers those early years. The CD-ROM version of the SCI goes back to the 80's but covers fewer publications.

In principle, all three databases can be used for the analysis of scientists' mobility, but all of them have also limitations. The *Web of Science* includes all scientific fields and can therefore be applied to all mobility studies in the sciences where most of the produced knowledge is published in the covered journals. The specialised databases *PubMed* and *INSPEC* have the advantage of containing fewer homonyms because they are restricted to a smaller number of fields. However, the problem of homonyms occurred in all databases, especially because none of the databases provides the full

names of the authors, but only initials. I could solve the homonym problem by using one or several of the following methods: I did a content analysis of the titles of the journal articles and in extreme cases also of the abstracts. If the online full text version of a journal article was available I checked the first name of the author. If the author had the same co-authors over a longer period, I used co-author names as context information. In a few cases I looked up self-citations of the author to get the publication track of one person.

PubMed and *INSPEC* only provide the first author's address. In order to use the database address information, it is necessary that the author is often enough the first author. A similar problem we have with the *Web of Science* by providing the corresponding authors address. The other addresses that are not directly linked to the author names can be used only as context information.

A disadvantage is that *INSPEC* and often also the *Web of Science* don't have the address information available for earlier years; *INSPEC* didn't collect it in earlier years, the *Web of Science* is often not available for earlier years. Thus, a more time-consuming bibliometric analysis by using the hard-cover versions of the journal articles must be applied. An alternative is to add the missing information by using non-bibliometric methods like extracting CV data from the internet.

Results

Empirical results

Due to the aim of the article, the main results are of a methodological nature. The empirical study of the Angiotensin scientists was used as a test example. The results of this study will shortly be presented. Of the 131 Angiotensin scientists investigated:

- 59 have always been in the USA;
- 34 moved to the USA, with
 - 18 of them still staying there, and
 - 16 having moved back to their home countries after a temporary stay;
- 3 moved from the USA to other countries;
- the remaining 35 elite members stayed in or moved to other countries than the USA.

In the case of the 18 scientists who moved into the USA and who are still staying there the time-span of their stay in the USA was checked: 12 stayed there more than

15 years, two ten years, one seven years and two about five years. Thus, the majority of them appears to have migrated permanently. These results confirm science policy's fear of a brain drain: A "drift" of elite scientists towards the USA appears to exist.

Methodical results

The methodical results are presented along with the three problems that had to be solved.

Delineating a specialty: The research field classification provided by ISI does not support a fine-grained delineation of scientific specialties. Using Gordon conference participant data has made it possible to identify a specialty via its elite. However, this method cannot be applied systematically because not all Gordon conferences are specialised enough to include only one specialty, and because not all specialties have Gordon conferences. Therefore, advanced word-based mapping methods are an alternative to be considered. However, detailed tests of these methods would be necessary before they could be used for delineating a specialty in terms of scientists rather than literature. For example, Lewison observed that his method doesn't recall articles in general journals without title keywords, which represent a large share of the literature (LEWISON, 1999: 533). Moreover, the lack of access to the ISI world data in a form that allows this kind of citation analysis often hinders the use of these methods.

Identifying a specialty's elite: ISI Essential Science Indicators could not be used for that purpose because of the described weaknesses of these indicators. Since they are applied on the individual level for only a small group of authors, these weaknesses cannot be treated as statistical error. Using participant lists of elite conferences, combined with a citation and co-citation analysis of the participants' publications, enabled at least an approximate solution to the problem.

Identifying international mobility: Bibliometric methods are in principle suitable to identify a scientist's international mobility. They appear to be the best solution because there are still not too many CV's available on the internet. All three databases tested were successfully used for this purpose. The following problems occurred:

- Not all addresses of all co-authors of a publication are available, only the first authors' address or the corresponding authors' address; there have to be enough publications where the focal scientist has main responsibility;
- Homonyms are a problem in all databases; and

- “Older” scientists’ first publication are not in the databases and must be searched in the hard-copy of publications or by the use of non-bibliometric methods.

Conclusions

In order to detect patterns of causation of the “brain drain”, we need really fine-grained studies of scientists’ interorganisational mobility. That means we have to go on the level of single specialties, even if the analysis starts with one country. The elite migration in science that triggers discussions about “brain drains” and “brain gains” has proven to be difficult to investigate. The more widely used concept of “scientific excellence” is too fuzzy to address the special role played by the elite of a scientific specialty. Elite members do not merely make the most important contributions to a specialty’s body of knowledge but simultaneously have an orienting function for all specialty members. From this follows that three conceptual and methodological problems must be solved in an investigation of elite migration: scientific specialties must be delineated, elite members must be identified, and the latter’s spatial mobility must be observed.

If we accept that the “value” of an elite member is partly due to his or her ability to orient a specialty, then an elite can be identified only in relation to a specialty. The first methodological problem is therefore the reliable delineation of specialties. This task has yet to be solved. Standard classifications of scientific fields as provided by ISI’s databases are too coarse to identify the relatively small communities of scientists whose work is organised by the reference to a shared body of knowledge.

Using participant lists of elite conferences has been successful in the case of the Angiotensin specialty. But the applicability of this approach is generally limited to certain specialties. Bibliometric methods based on co-citations or occurrence of words suggest themselves. However, there are still major obstacles to be overcome. Firstly, both co-citation clustering and methods based on frequencies of word occurrences apply arbitrary thresholds. Secondly, co-citation clustering and bibliographic coupling seem to recall only part of a specialty’s literature. Thirdly, the word-based methods’ reliance on expert judgements necessarily introduces these experts’ perspective on the field. Thus, the problem of delineating specialties is still one of the major challenges to scientometrics (VAN RAAN, 1997: 215). However, bibliometric methods as a general approach for delineating specialties seem to be the most promising direction for development. In any case, the delineation of scientific specialties will always have to take into account an uncertainty that is caused by the fuzzy nature of the specialty itself.

If elite members are characterized by their orienting function for a specialty, methods for identifying an elite have to address this relation. Thus, bibliometric

methods that are based on citations cannot simply count all citations a scientist gets but must identify the citations a scientist gets within his or her specialty. This approach again depends on solving the problem of specialty delineation. The combination of elite conference participation and citation analysis within the sample of active conference participants was only a provisional approach that cannot be applied to all specialties.

The conventional methods for the measurement of elite scientists' mobility have several shortcomings that prevent a systematic application. The use of bibliographic databases is an alternative that can overcome these shortcomings. The address information provided by bibliographic databases can be used to track authors' spatial mobility. However, the use of bibliographic databases is not a quick and easy method. The data has to be purified from homonyms and sometimes must be supplemented with data from the full-text versions of publications or from non-bibliographic sources. This additional work is minor in fields with a high publication output per year.

The results of my test study of the biomedical specialty Angiotensin confirmed the supposed brain drain towards the United States. Further studies of other specialties are needed to find out field-specific patterns of causation. For example, the existence and strength of a Matthew effect on the national or organisational level should be investigated. For the science policy measures of "buying back" scientists to be successful, they must encompass not only the rather obvious attractors such as competitive salaries and excellent working conditions, but also the possibly hidden attractors such as presence of other elite members, opportunities for interdisciplinary collaborations etc. that may be at work but can be assumed to vary between specialties.

*

Special thanks to Beverley Biglia for her assistance in the search for CV data from the internet and in testing the suitability of the three databases for mobility measurements. Jochen Gläser rescued me from getting stuck in some methodological problems by providing conceptual input and spending many hours of programming Microsoft VBA Macros. I thank him and Linda Butler for critical reading the manuscript.

Certain Data herein are derived from the Australian National Citation Report prepared by the Institute of Scientific Information, Inc. (ISI), Philadelphia, Pennsylvania, USA.

References

- AKSNES, D. W., T. B. OLSEN, P. O. SEGLEN (2000), Validation of bibliometric indicators in the field of microbiology: A Norwegian case study. *Scientometrics*, 49 : 7–22.
- ALLISON, P. D., J. S. LONG (1987), Interuniversity mobility of academic scientists. *American Sociological Review*, 52 : 643–652.
- AMICK, D. J. (1974), An index of scientific elitism and the scientist's mission. *Science Studies*, 4 : 1–16.

- AUSTRALIAN RESEARCH COUNCIL (2001), Federation Fellowships. *Discovery*, December: 1-2 (<http://www.arc.gov.au/newsletter/dec01/page1.htm>).
- BMBF (2001), *Beitrag des BMBF zum Fragenkatalog für die öffentliche Anhörung "Chancen und Risiken der Informationsgesellschaft" am 28. Mai 2001*. (Kdrs. 14/10b). Berlin: Deutscher Bundestag, (http://www.bundestag.de/gremien/welt/weltto/weltto121_stell007.pdf).
- BUECHTEMANN, C. F. (Ed.) (2001), *Deutsche Nachwuchswissenschaftler in den U.S.A. - Ergebnisse der Vorstudie*. Bonn: BMBF.
- CRAWFORD, S. (1971), Informal communication among scientists in Sleep Research. *Journal of the American Society for Information Science*, 22 : 301-310.
- DEBACKERE, K., M. A. RAPPA (1995), Scientists at major and minor universities: Mobility along the prestige continuum. *Research Policy*, 24 : 137-150.
- DIETZ, J. S., I. CHOMPALOV, B. BOZEMAN, E. O'NEIL LANE, J. W. PARK (2000), Using the curriculum vita to study the career paths of scientists and engineers: An exploratory assessment. *Scientometrics*, 49 : 419-442.
- GLÄSER, J. (2001), *Scientific Specialties as the (Currently Missing) Link between Scientometrics and the Sociology of Science*. Paper presented at the 8th International Conference on Scientometrics & Informetrics, Sydney, Australia, July 16-20th.
- GRAVERSEN, E. K., S. O. NÅS, A. EKELAND, M. BUGGE, C. SVANFELDT, M. ÅKERBLOM (2002), *Knowledge Transfer by Labour Mobility in the Nordic Countries*. Working papers 2002/1. Aarhus: Analyseinstitut for Forskning.
- HELLEMANS, A. (2001), Beating the brain drain. *Naturejobs*, 414 : 4-5.
- LAUDEL, G. (1999), *Interdisziplinäre Forschungsk Kooperation: Erfolgsbedingungen der Institution "Sonderforschungsbereich"*. Berlin, Edition Sigma.
- LAUDEL, G. (2001), Collaboration, creativity and rewards: why and how scientists collaborate. *International Journal of Technology Management*, 22 : 762-781.
- LEVIN, S. G., P. E. STEPHAN (1999), Are the foreign born a source of strength for US science? *Science*, 285 : 1213-1214.
- LEWISON, G. (1999), The definition and calibration of biomedical subfields. *Scientometrics*, 46 : 529-537.
- MAHROUM, S. (1999), Global magnets: Science and technology disciplines and departments in the United Kingdom. *Minerva*, 37 : 379-390.
- MARTÍN, M.-J., L. M. PLAZA, J. REY (1996), Transnational mobility of researchers: the Spanish Fellowships Programme for "senior" scientists. *Research Evaluation*, 6 : 13-18.
- NOYONS, E. C. M., A. F. J. VAN RAAN (1995), *Mapping the Development of Neural Network Research: Structuring the Dynamics of Neural Network Research and an Estimation of the German Activity*. Report to the German Federal Ministry of Education and Science (BMBF) CWTS 95-06. Centre for Science and Technology Studies, Leiden University.
- NOYONS, E. C. M., A. F. J. VAN RAAN (1998), Monitoring scientific developments from a dynamic perspective: Self-organized structuring to map neural network research. *Journal of the American Society for Information Science*, 49 : 68-81.
- NOYONS, E. C. M., R. K. BUTER, A. F. J. VAN RAAN, H. SCHWECHHEIMER, M. WINTERHAGER, P. WEINGART (2000), *The Role of Europe in World-Wide Science and Technology: Monitoring and Evaluation in a Context of Global Competition. Report for the European Commission*. Leiden: Universiteit Leiden, http://www.uni-bielefeld.de/iwt/mw/allgemeines/tser_bericht.pdf.
- NOYONS, E. (2001), Bibliometric mapping of science in a science policy context. *Scientometrics*, 50 : 83-98.
- NOYONS, E. C. M., R. K. BUTER, A. F. J. VAN RAAN (2002), Field delineation for bibliometric evaluation. Paper presented at the *Seventh International S&T Indicators Conference "The Challenge of Measuring Knowledge Structures in Science and Technology"*, Karlsruhe, Germany, 25-28 September.
- PICKRELL, J. (2001), Fighting brain drain - Ireland gives its stars a big pot o' gold. *Science*, 293 : 1028-1029.

- PIERSON, A. S., P. COTGREAVE (2000), Citation figures suggest that the UK brain drain is a genuine problem. *Nature*, 407 : 13.
- ROSENFELD, R., J. A. JONES (1987), Patterns and effects of geographic mobility for academic women and men. *Journal of Higher Education*, 58 : 493–515.
- SCOTT, J. (1991), *Social Network Analysis: A Handbook*. London. London, Sage.
- SHAUMAN, K. A., XIE YU (1996), Geographic mobility of scientists: Sex differences and family constraints'. *Demography*, 33 : 455–468.
- SINCELL, M. (2000) Why do the world's leading scientists flock to the United States? *Physics World*, 13 : 10–11.
- SMALL, H. G. (1977), A co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social Studies of Science*, 7 : 139–166.
- SÖDERQVIST, T., A. M. SILVERSTEIN (1994), Participation in scientific meetings: A new prosopographical approach to the disciplinary history of science. The case of immunology, 1951-72. *Social Studies of Science*, 24 : 513–548.
- SPURGEON, D. (2000), Canada tries to limit nanotech brain drain. *Nature*, 408 : 623.
- STEPHAN, P. E., S. G. LEVIN (2001), Exceptional contributions to US science by the foreign-born and foreign-educated. *Population Research and Policy Review*, 20 : 59–79.
- STIFTERVERBAND FÜR DIE DEUTSCHE WISSENSCHAFT (2002), *Brain drain - brain gain: Eine Untersuchung über internationale Berufskarrieren*. Kassel, GES.
- TIJSEN, R. J. W., M. S. VISSER, T. N. VAN LEEUWEN (2002), Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics*, 54 : 381–397.
- TRIMBLE, V. (2000), Some characteristics of young vs. established American astronomers: Entering the new century. *Scientometrics*, 48 : 403–411.
- VAN HEERINGEN, A., P. A. DIKWEL (1987), The relationships between age, mobility and scientific productivity. Part I. Effect of mobility on productivity. *Scientometrics*, 11 : 267–280.
- VAN RAAN, A. F. J. (1997), Scientometrics: State-of-the-art. *Scientometrics*, 38 : 205–218.
- VAN RAAN, A. F. J. (2000), The Pandora's box of citation analysis: Measuring scientific excellence – The last evil? B. CRONIN, H. B. ATKINS (Eds), *The Web of Knowledge*. Medford, New Jersey: Information Today, Inc., pp. 301–319.