Net ReAct

The role of networking in research activities

Deliverable 1.1

Characterisation of doctoral students

by

Franz Barjak

University of Applied Sciences Solothurn Northwestern Switzerland

Deliverable Version:	deliverable 1.1, v3-2_korr.doc, final draft
Delivery Date:	10.05.2005
Contract:	The role of networking in research activities EU Contract No. 22540-2004-12 F1ED SEV DE
Issued by:	Institute for Prospective Technological Studies, Commission of the European Communities
Contractor:	empirica GmbH Oxfordstr. 2 D - 53111 Bonn e-mail geschäftsführung@empirica.com phone +49 (0)228-98530-0 fax +49 (0)228-98530-12

empirica

University of Applied Sciences Solothurn Northwestern Switzerland
Engineering - Business - Social Sciences

Table of contents

Executive Summary 5
1 Introduction to the NetReAct project and the deliverable 12
1.1 Overall aim and structure of the project 12
1.2 Contents of this deliverable 13
2 The role of networking activities in public research 15
2.1 Research teams as the nucleus of scientific knowledge production 15
2.1.1 Research teams and research networks – two interlaced structures in science 15
2.1.2 Composition of research groups 18
2.1.3 Further group-related influences on research productivity 23
2.2 Research collaboration and productivity 26
2.2.1 What is research collaboration? 27
2.2.2 The impact of research collaborations on research productivity 27
2.3 Further environmental covariates of research productivity 33
2.4 Summary of the literature review, hypotheses and operationalisation of the hypotheses 35
2.4.1 Measuring research productivity 35
2.4.2 Research teams and their influence on research productivity 38
2.4.3 Research collaboration and research productivity 44
2.4.4 Further covariates of research productivity 46
3 Research approach in work-package 1 48
3.1 General overview of the NetReAct approaches 48
3.2 Survey 49
3.2.1 The research population 49
3.2.2 Sampling and sample characteristics 51
3.2.3 Survey methodology 53
3.3 Exploratory interviews 53
4 Interview results in regard to research teams and doctoral students in the life sciences 53
4.1 Research teams 53

4.1.1 Issues related to time and discipline of the research team 53
4.1.2 Personnel structure and development 53
4.1.3 Research funding 53
4.2 Doctoral students 53
4.2.1 Recruitment of doctoral students 53
4.2.2 Fulfilment of the implicit contract 53
4.2.3 Demographic and education-related information 53
5 Outlook for further work to be done in WP 1 53
References 53
Annex 53

Preface and acknowledgement

This deliverable was mainly written by Franz Barjak, University of Applied Sciences Solothurn Northwestern Switzerland (FHSO). However, it used several inputs contributed from the other NetReAct partners which should be acknowledged:

- The identification of the research population and the sampling was carried out jointly by Mike Thelwall and Xuemei Li from SCIT and Franz Barjak from FHSO; input for Hungary was obtained from Andras Schubert, IRO HAS Budapest; for the Czech Republic from Irena Bartova, SC&C Prag, and for Spain from Mercedes Argüello Casteleiro, University of Santiago de Compostela, Spain.
- The address collection for the survey was done at FHSO with the help of several research assistants: Antoinette Berta, Michele Bolzern, Stefanie Strauss, Thomas Stuber, Daniel Sumi, and Kevin Ziegler.
- The exploratory interviews were carried out by Tobias Hüsing (empirica), Mike Thelwall and Alesia Zuccala (both SCIT), and Franz Barjak (FHSO).
- Small parts of the deliverable text draw heavily on previous NetReAct documents written by different partners of project, in particular Mike Thelwall (SCIT), Wolfgang Glänzel (KU Leuven), Simon Robinson (empirica), and Franz Barjak (FHSO).
- All the NetReAct partners read this deliverable and improved it through providing comments and corrections.

Executive Summary

The present report provides an overview of the work done so far in the work-package 1 "Characterisation of doctoral students". It contains three main parts:

(1) The review of the scientific literature on the role of doctoral students, post-doctoral researchers and research collaborations in scientific knowledge production, focusing whenever possible on the life sciences (chapter 2).
(2) The description of the methods and results of the sampling of life sciences research teams in the 10 NetReAct countries and the current structures of the research population and the sample (chapter 3).
(3) The interpretation of the results of exploratory interviews carried out with team leaders in the life sciences. The interviews targeted all different themes of the project. However, the interpretation in this report is limited to the questions on the research teams and doctoral students (chapter 4).

The other chapters contain a brief introduction (chapter 1) and an outlook for the further work to be done in this work-package (chapter 5).

The role of networking activities in public research – summary of the literature review and hypotheses

In the literature review and also in the first empirical steps of the project we encountered one key problem that develops around the lack of a clear micro-level unit of analysis. Even if public research outside of universities is excluded, the organisation structure in public science still varies considerably across countries and even within countries (e.g. between regions or universities). Therefore it is nearly impossible to identify a unit of analysis that has the same meaning everywhere. We chose "research teams" which are the smallest organisation unit above the individual scientist, as these seem to be present in all countries. Previous studies have delimited these teams in different ways, such as taking co-authors, scientists with the same institutional affiliation, or scientists working together. In order to be able to investigate the main questions of the project but at the same time reduce the danger of misunderstandings in the data collection, we chose a definition of research teams that is based on joint work. Hence we define a research team as a group of people, scientists and non-scientists, which work in the same location for a longer time period to produce new scientific knowledge. Research teams are linked – mainly through their team leaders or principal investigators (PI) – to other research groups in what we and others before us have called "research networks". The research networks constitute the "platform" on which the interactions (e. g. research collaborations, personnel exchanges, knowledge exchanges) between research groups take place.

The composition of research teams is one of the key issues of interest of the present project. The literature points to several elements of the team composition which might be related to research productivity: the presence of highly productive scientists, the number of post-doctoral researchers (who are estimated to make up between 9 and 19% of all team members), the presence of other personnel groups (permanent scientific staff, technicians and administrators, guests and visitors), an interdisciplinary team composition, and differing national and cultural backgrounds. However, the available empirical evidence on these issues and their relationship to research productivity is scarce. PhD students are estimated to constitute on average 30% of the team members (in the life sciences as well as in other disciplines). Though several positive influences of PhD students on productivity have been proposed, too, the empirical research has failed to corroborate these so far. One explanation is that PhD students only produce a part of the needed ingredients to a good team performance, namely the experimental results, and that principal investigators still have to add their contribution, mainly the interpretation and link-up to the state-of-the-art in the field. Another explanation might lie in breaches of the implicit contract between PhD students and their supervisors.

The second key issue of interest of this study are research collaborations between teams. Research collaborations can have a range of forms and purposes. We focus on collaborations in which all involved groups do creative work and which consequently can be identified through the joint authorship of publications. Among the motivations for research collaborations we can distinguish between cognitive, financial, social and non-research motivations which all influence the design of collaborations and probably also their results. In most empirical studies, research collaborations correlate positively with performance and productivity measures of scientists and research teams. However, at individual level a curvilinear relationship has been found: up to a certain size of the collaboration network, scientists publish more, the more they collaborate; beyond this threshold, however, the effect of additional collaboration partners is reversed and negative. Certain features of collaborations seem to have exceptionally positive impacts on research productivity, namely the involvement of international partners and the collaboration of public research organisations with private partners. Though both also entail costs, the benefits seem to prevail.

There are several other group-related factors, e.g. group size and age, resources of the group and group interactions, and factors related to the parent organisation, the scientific discipline and the country/science system which affect research productivity, too. These factors should also be taken into account in an empirical analysis.

The theoretical considerations and empirical findings can be condensed into a range of hypotheses which will guide the empirical work:

> H1.1: The structure of the research teams led by highly productive scientists is more diverse than the structure of teams led by less productive scientists.
>
> H1.2: The collaboration partners of research teams led by highly productive scientists are more diverse than the collaboration partners of teams led by less productive scientists.
>
> H1.3: The research productivity of teams led by highly productive scientists is higher than the research productivity of teams led by less productive scientists.

H2.1: Research teams that realise a division of labour between junior and senior scientists have higher research productivity than teams that don't realise a division of labour.

H2.2: Research teams with many PhD students do not have higher research productivity than teams with few PhD students.

H2.3: The research productivity of teams fulfilling the implicit contract with PhD students is higher than the research productivity of teams that break it.

H3.1: Research teams with many postdoctoral researchers have higher research productivity than teams with few postdoctoral researchers.

H3.2: The research productivity of teams fulfilling the implicit contract with postdoctoral researchers is higher than the research productivity of teams that break it.

H4.1: Research teams with PhD students from other disciplines are more productive than research teams without such students.

H4.2: Research teams with postdoctoral researchers from other disciplines are more productive than research teams without such postdoctoral researchers.

H5.1: Research teams with PhD students from other countries of birth are more productive than research teams without such students.

H5.2: Research teams with PhD students who obtained their graduate degree in another country are more productive than research teams without such students.

H5.3: Research teams with postdoctoral researchers from other countries of birth are more productive than research teams without such postdoctoral researchers.

H5.4: Research teams with postdoctoral researchers who obtained their PhD and/or graduate degree in another country are more productive than research teams without such postdoctoral researchers.

H6.1: Collaborating research teams have higher research productivity than non-collaborating teams.

H6.2: The marginal effect of additional research collaboration on research productivity is decreasing (curvilinear relationship).

H7.1: If research teams predominantly collaborate for cognitive purposes, then their collaborations will have a positive effect on their research productivity.

H8.1: Research teams which collaborate internationally have higher research productivity than teams which do not collaborate internationally.

H8.2: Research teams which collaborate with research teams in the US, Switzerland and the UK have higher research productivity than teams which do not collaborate with teams from these countries.

> H8.3: Research teams involved in interdisciplinary research collaborations have higher research productivity than teams which do not collaborate with teams from other disciplines, provided that they have the same team structure.
>
> H8.4: Research teams involved in research collaborations with private partners have higher research productivity than teams which do not collaborate.
>
> H8.5: Research teams involved in research collaborations with private partners have lower research productivity than teams which collaborate with public research organisations.

In addition to the hypotheses, the summary section of the literature review also includes a first attempt to operationalise the concepts. The analysis of the hypotheses will be carried out mainly in work-package 4 and it will use all the data sets which will be assembled within the NetReAct project. Work-package 1 also identified the overall research population of the analysis and drew the sample of research teams.

The research population and sample

The primary statistical unit of analysis is the university-based research team. Through the examination of the International Handbook of Universities and university websites we identified more than 6,300 research teams in the life sciences delimited according to ISCED 1997. The teams are affiliated to 356 universities in the 10 European countries included in the analysis. On average there are 18 teams per university.

From this research population of 6,300 research teams we drew a sample of 1,500 teams which will be included in the survey. The sampling drew on information from the internet and webometric methods. First, strata were built according to country and the number of hyperlinks pointing to the teams' websites, as previous research has shown that the number of hyperlinks is correlated with scientific performance. Then, drawing probabilities for each stratum were calculated which were inversely related to the hyperlink metric. Third, from each stratum the number of teams corresponding to the drawing probability was randomly selected. This procedure produced a sample of research teams stratified by country and the hyperlink metric. The sample includes more than 200 teams in the countries with large life sciences research systems (Germany, France, and the UK), more than 125 teams in the countries with medium-sized life sciences research systems (Spain, Italy, and Sweden), and more than 100 teams in the countries with small life sciences research systems (Czech Republic, Hungary, Norway, and Portugal). Between 69% (Portugal) and 16% (France) of all research teams in a country are included in the sample. On average only 5 teams per university could be included (maximum: 21 teams in Norway, minimum: 3 teams in Italy and the UK). This will have some consequences for the analysis at university level: it will only be possible for the few larger universities in each country (depending also on the response rate), but not for all 302 universities included in the sample.

On average the teams in the sample received 8.5 hyperlinks pointing to their website. One fifth received more than 10 links and one third no links at all. Three out of a hundred teams don't even have a website. All these statistics support the finding that teams from Southern European countries and the new member states tend to be less visible on the internet than teams from the UK, Germany and the Nordic countries.

As we intend to carry out the survey via e-mail, the names and e-mail addresses of the team leaders had to be collected, too. Within this process of address collection, additional data on the teams was also obtained which will serve the purpose of checking the representativeness of the responses to be obtained in the survey. The average size of the research teams is 18.4 scientific and non-scientific workers, ranging from 13.4 in Portugal to 22 in the UK and Hungary. Across all countries PhD students represent 32.5% of the total scientific and non-scientific staff which is in the range of the figures obtained in previous studies. PhD students are particularly relevant in Portuguese, British, Czech, and Spanish teams, whereas they are a lot less present in Hungarian and Norwegian teams. Post-docs constitute nearly one fourth of the personnel. According to this first data collection, they are more important in the UK (32%), Portugal (30%), Italy (28%), and Norway (27%) and less important in France (9%) and Hungary (13%). However, it has to be noted that the figures on PhD students and post-docs are very much preliminary and only suitable for getting a vague overview of the sample structure. They were gathered from the World Wide Web, and even for those teams which provided them, we do not know, whether they are up-to-date and complete (probably in many cases they are not).

In order to test the feasibility of certain settings of the survey, gather background information on research in the life sciences, and test different question approaches and formulations, exploratory interviews with the heads of research teams in the life sciences were carried out. We interviewed the heads of four teams in France, three teams in the UK and two teams in Germany.

Interview results in regard to research teams and doctoral students in the life sciences

A major issue of the interviews was to evaluate the retrospective data collection for the year 2003 which is specified in the tender and necessary to obtain correspondence with the publication data. Most of the interviewees were able to answer the questions for the year 2003. The main difficulty lay in the allocation of developments and trends to time periods and in the description of the situation before 2003. This will be taken into account in the survey by clearly specifying the time period whenever information on developments is required.

The tender specifications request the use of the ISCED 1997 classification for defining the teams' scientific discipline. ISCED 1997 is oriented towards education and not towards research. The interviewees who were asked to classify their team according to ISCED 1997 struggled with this task: they usually answered the question from the standpoint of their research products, listed several disciplines and settled on one or two general ones, and stated that the question was difficult to answer as their main field was included in various possible responses. The graduate education and training activities of the life sciences teams are very much connected to their research activities. Therefore, we think it would be better to use a research-related classification instead of ISCED 1997. This should reduce the difficulties of the respondents to provide an appropriate classification of the team and it raises the validity and reliability of this information. Also, the combination of the survey and bibliometric datasets will be easier in the data analysis. For these reasons, we propose to use the K.U. Leuven – IRO Subject Classification only.

The research teams of the interviewees had between 10 and 30 members. In most cases there are more PhD students than post-docs. The interviews could not solve the question whether master students should be included among the team members as the respondents themselves had differing opinions on this. However, without a doubt guests and visitors should be left out, as they are normally drawn to a team to learn specific techniques and methods, and do not enhance the team's research capacities themselves. Out of the nine interviewees three stated that they didn't or couldn't pursue any strategies or policies in regard to the composition of their research team. Three respondents said that they would look for certain qualifications in regard to the field of education or knowledge of theories, and two interviewees highlighted that they would appreciate a mixture of qualifications and nationalities in their teams. Out of the nine teams interviewed, five had grown, two had remained constant and two had decreased in the previous five years. Growth was often realised through acquiring research grants and recruiting more post-docs and PhD students.

The recruitment of PhD students seems to follow one of three models: In the "graduate school model" graduate students are recruited in a national or even international competition organised by a committee of principal investigators. At the beginning of the "job competition model" usually stands a grant or studentship that creates an open position which the team leader fills with his own master graduates, through selecting an applicant from a pool of unsolicited applications, or through advertising the job opportunity. Other team members are often involved in the recruitment process. The most important factors in job competitions which are part of the two proposed models are the qualifications and technical skills, motivation, and social competence of the applicants. They are assessed in job interviews and by means of evaluating the past performance of the applicants. The "application model" is similar to the job competition model, but vice versa: first, the team leader receives an interesting application, and then he has to look for funding or help the student to-be to obtain a studentship. Whether the models are related to research productivity remains an open issue.

The literature on doctoral students stresses the importance of the so-called "implicit contract" between PhD students and supervisors. The implicit contract means that the doctoral students with their experimental work contribute to research programmes from which the supervisor draws reputational benefits. The students are not immediately compensated for this work, but they receive the compensation after finishing their PhD thesis, when the supervisor and the projects they did under his guidance help them to obtain follow-up positions. In recent analyses of the job market of PhDs for the US and France we find the concern that the implicit contract is increasingly broken, when PhD supervisors cannot support the further career of their students anymore. This, in turn, reduces the attractiveness of a team for PhD students. In the interviews we investigated this in three ways: by asking the interviewees for the functions of their PhD students, looking for practices of exploiting PhD students, and getting information on the positions that PhD graduates obtained. Both, the questions on the functions and practices of exploiting PhD students are subject to specific problems, and consequently they will be dropped from the questionnaire. The question on the positions of graduates after receiving their PhDs, however, was readily answered by most respondents. And it seems to be useful to assess the success of the PhD education.

Outlook for further work to be done in WP 1

The further work in work-package 1 centres on collection and analysis of data on doctoral students and the coordination with the other work-packages of the project. The first task involves several steps: (1) Refinement of the sample and finalisation of the address collection; (2) development of the questionnaire; (3) pretesting the questionnaire; (4) realisation of the survey; (5) data analysis; (6) methodological generalisation. The second task mainly targets at the interdependence between the different work-packages. It secures that the data obtained in the work-packages 1-3 can be integrated in a joint analysis in work-package 4.

1 Introduction to the NetReAct project and the deliverable

1.1 Overall aim and structure of the project

The "The role of Networking in research activities" (NetReAct) project is designed to capture, describe and analyse the strategies, patterns, dynamics and impact of networking in research activities in the life sciences in 10 European countries. The objectives are as in the Tender Specifications, specifically to develop and apply methods to shed light on current research collaboration behaviour of European universities, in particular in respect of the mobility of ideas and of personnel (brains) and in doing so:

- to assess the dynamics of universities' networking activities in respect of other universities, of public and private research bodies and
- to measure the capacity of universities and their laboratories to attract doctoral students and post-doctoral staff from other geographical areas.

These objectives will be pursued with different methodological approaches of data collection (and interpretation) and an analysis that integrates the different data sets. The work is structured along four work-packages:

- Work-package 1: Characterisation of doctoral students
- Work-package 2: Bibliometric measures of networking activities and of their impact
- Work-package 3: Characterisation of post-doctoral training
- Work-package 4: Analysis of the datasets

Work-package 2 is split up into two tasks:

- Task 2.1 Measures of networking activities and of their impact based on the ISI databases
- Task 2.2: Measures of networking activities and of their impact based on the internet

The main unit of analysis in the present study is the research team – we use the term research group as synonym (for an operational definition see chapter 2.4.2). This is due to several reasons:

- The heterogeneous construction of science systems across countries makes an analysis at the levels of universities or departments difficult, as the terms might stand for different types of organizations; this problem can be reduced, if the analysis goes down to the micro level of groups and collaborations (van Raan, 2004).
- The evidence of previous research suggests that the productivity of individual researchers is shaped mainly by the smaller teams and collaborations with which they are most closely involved (von Tunzelmann, Ranga, Martin, & Geuna, 2003).

- At the level of research teams complementarities between different types of personnel, reputation externalities, and scale effects can be explored (Carayol & Matt, 2004).
- Last but not least, the survey-based research approach required that we chose a unit of analysis, for which respondents could answer detailed (and retrospective!) questions on the characteristics of the doctoral students and postdoctorants in the team. This was not deemed possible for larger units of analysis such as entire departments or schools.

1.2 Contents of this deliverable

The NetReAct deliverable D1.1 summarises the work done during the first four project months mainly in the work-packages 1 and 3. It rests on three pillars of work which are the basis for the following chapters 2-5:

1) Review of scientific literature on the factors explaining the research productivity of research groups with a particular focus on group structure and collaboration between groups,
2) Identification of life sciences research teams in the 10 NetReAct countries mainly from the internet,
3) Realisation of exploratory interviews with selected scholars and research group leaders from the life sciences.

Chapter 2 reviews the previous empirical work on the research productivity of research groups. It starts with a discussion of the configuration of research teams and the resulting consequences for research productivity. In particular, the literature on highly productive scientists, doctoral students and young researchers, and post-doctoral researchers is reviewed. Moreover, we present findings on interdisciplinary and multinational research teams and further, team-related aspects, such as team size, material resources, and group interactions. Chapter 2 also includes a survey of the literature on research collaborations, or more specifically on the role of research collaborations for research productivity. The structure of research teams and their collaboration with other teams are not the only variables which have some bearing on the output of research teams. Empirical analyses should take further covariates into account in order to prevent spurious correlations: the characteristics of the parent organisation, scientific discipline, and country are three such environmental parameters which should be included in an analysis, as section 2.3 shows. The final part of chapter 2, section 2.4, summarises the main results of chapter 2 and formulates hypotheses which guide the empirical work of work-packages 1, 3, and 4, in particular. Of course, these hypotheses are still intermediary, and the further work will lead to restatements and modifications.

In chapter 3 we provide a description of the empirical work of work-packages 1 and 3 that has been done so far. In particular, we explain how the research population of the survey was identified and how the sampling was carried out. Furthermore, selected characteristics of the sample are described and the survey methodology is briefly presented. The survey is the main empirical pillar of work-packages 1 and 3. Additional exploratory interviews are needed to ground the survey instrument in the actual research practices in the life sciences. The interviews should also ensure that the survey questions are answerable and as comprehensive as possible, and that misunderstandings are avoided. Chapter 3 elaborates briefly on the methodological approach to these interviews, too.

Chapter 4 presents the main results of the interviews in regard to the structure of research teams and the role of doctoral students in research teams. In regard to the latter issue, the recruitment strategies employed for filling open positions for doctoral students and the implicit contracts which govern the work of doctoral students are discussed specifically.

At this stage of the research it is not possible to draw conclusions. The last chapter therefore gives an outlook on the further work that will be done in work-package 1 and how this interlinks with the other work-packages of the NetReAct project.

2 The role of networking activities in public research

The present chapter presents an overview of the state of the art on scientific research teams and their research productivity. It starts with a discussion of research teams and how their internal structure, size, resources, and culture affect research productivity. Then, particular emphasis is put on the role of collaboration for research productivity (2.2). A third section deals with further covariates of research productivity: the characteristics of the parent organisation, the scientific discipline and geographical location (country). The fourth section (2.4) summarises the literature review, deduces hypotheses and presents a first attempt at their operationalisation.

Parts of this chapter will be included in the final WP 1 report and other parts are further discussed and improved in the WP 4 deliverable.

2.1 Research teams as the nucleus of scientific knowledge production

2.1.1 Research teams and research networks – two interlaced structures in science

Research teams

Research teams (or groups) and research networks (or scientific communities, invisible colleges) were considered as the basic social entities in the sciences since the early works on the social organisation of science (Hagstrom, 1965; Price & Beaver, 1965). They are the basic production unit of scientific knowledge in which scientists work with students and technicians which receive some sort of compensation, e.g. recognition, training, monetary compensation (Hagstrom, 1965). They are groups bound by solidarity, forged to recruit and socialize new members in a research area, secure the continuation of research, and increase the production by "squeezing papers out of the rather large population of people who have less than a whole paper in them" (Price & Beaver, 1965, p. 1015).

The identification of research teams is not a trivial issue as no clear institutionalised delimitation exists. Analyses of the research productivity at this level use different approaches for defining research groups. The definitions also tend to influence the results and each has its strengths and weaknesses. In the literature three alternatives of delimiting research teams are proposed:

- The institutional definition takes all researchers in the same unit of an organisation (Cohen, 1981; Hagstrom 1965). For this definition it has to be decided, how technicians and administrative staff, PhD students, guests and visitors are treated.
- The functional definition of research teams is based on joint research activities. It can combine several different organisational forms and result in teams limited to one single organisation or teams drawing membership from several separate

organisations (UNESCO study see Andrews, 1979). The functional definition has to specify what joint research actually means – the UNESCO research team, for instance, specified it temporally with a minimum of eight hours of team-related work per week and scientist and a minimum group size of three scientists.

- The definition of research groups based on co-authorships (e.g. Bordons & Zulueta, 1997; Seglen & Aksnes, 2000) has to find methods and thresholds for adding co-authors to a particular group. It does not include unproductive scientists, and is therefore somewhat problematic when it comes to assessing the relationship between group size and research productivity (Stankiewicz, 1979).

Each of these definitions has its roots in the fact that scientists work together in the production of new knowledge. They differ in regard to how they assess this joint work and whom they count as a scientific worker.

In addition to research teams, further types of organisations that group together scientists and research teams and structure universities or other public research organisations exist. Their names, such as department, institute, lab, unit, centre etc., have varying meanings and connotations for different organisations and countries. Previous studies of these organisations used the term "laboratory" or "lab" to denote them (e.g. Joly & Mangematin, 1996; Laredo, 1999, 2001). Sometimes a team might be equalled with such a laboratory, in other cases two, three or more teams might add up to one. As Laredo (1999) shows, there exist significant differences across countries in this respect, with Spanish biomedical labs having rather few and French labs having many teams. Research networks are another social structure in science which is even less clearly delimited than research teams or labs but nevertheless exerts an important influence on the production of new scientific knowledge.

Research networks

Research networks or invisible colleges consist of the principal investigators and heads of research teams which work on a specific group of problems and topics in a research field. Mulkay (1977) calls them "elite members" or the "most productive and influential scientists". The informal network that connects them has been described as the "power group of everybody who is really somebody in a field" (Price & Beaver, 1965, p. 1011). Mulkay et al. 1975 define "scientific research networks" as relative intensive concentrations of interest ties which are characterized by a clustering of communication choices and which facilitate the development of a scientific consensus (cited in Lievrouw, 1988, p. 12). Any established scientist can be a member of different networks which can overlap and cross the boundaries of fields and disciplines (Mulkay, 1977). However, invisible colleges are not static but they show considerable dynamics of their membership depending on the in- and out-migration of scientists. Also, they are segmented into separate groups which are centred on the institution or working area of its key member(s) (Price & Beaver, 1965).

Invisible colleges control a significant share of the available resources for the work in a problem area and its members make most of the decisions on research strategies in the field concerned (Price & Beaver, 1965). As fields and in particular scientific disciplines have become too big and heterogeneous to permit an effective communication of research results, the networks constitute a filter that harmonises the amount of information with researcher's absorptive capacities and time frames for communication (Mulkay, 1977). They were found to serve as a channel for the diffusion of research ideas and research results: there is evidence that scientists outside or at the fringe of a research network received less information than its participants (Cole & Cole, 1973; Crane, 1972; Cronin, 1982; Mulkay, 1977). In some cases, research networks function as bourses for skills and for access to costly research apparatus (Mulkay, Gilbert, & Woolgar, 1975). Knorr-Cetina (1981, 1983) pointed to the "transscientific" or "transepistemic" nature of the networks. Transepistemic relationships include different groups of people that interfere in knowledge production, such as the researchers of various specialties, administrators, representatives of funding agencies or enterprises, publishers etc. which switch roles and functions.

Thus early studies have shown how research networks allocate resources in science, steer the direction of research, and govern the acceptance and diffusion of new knowledge. Research networks are the "platforms" on which interactions between the "nodes", individual scientists or research groups take place. Many different types of interactions exist between the network nodes, for instance joint membership in working groups of scholarly associations, professional committees, funding boards, or editorial boards of scientific journals; communication at conferences; or joint work on particular research projects (to be discussed in more detail in section 2.2).

Developments during the last 25 years – such as the growing need for scientific knowledge outside of science, the growing demand for a larger economic and social return from scientific work, the increasing financial squeeze of science funding, and the spread of new information and communication technologies – have to some extent changed the structures of research networks and the ways they operate, but there is no evidence to date that their importance has lessened. Several more recent explorations of "Post-Normal Science" (Funtowicz & Ravetz, 1993), "Mode 2" (Gibbons, Limoges, Nowotny, Schwartzman, Scott, & Trow, 1994), "Triple Helix" (Etzkowitz & Leydesdorff, 1997, 2000), "Techno-scientific communities" (Meyer-Krahmer & Schmoch, 1998), and "Knowledge-intensive communities" (David & Foray, 2002) point to a growing interaction between science and stakeholders from other areas of society, in particular from industry, politics and social interest groups. However, empirical studies tend to show that scientific and technical networks have some overlaps, but in general they stay apart (Debackere & Rappa, 1994; Murray, 2002; Rappa & Debackere 1992; Widhalm, Topolnik, Kopcsa, Schiebel, & Weber, 2001).

This applies in particular to the life sciences. Though they are linked in many ways to the biotechnological industry, the networks are not necessarily integrated. Several studies have shown that biotechnology firms depend to a large extent on scientific knowledge and give it a high priority among their resources (Zucker, Darby, Brewer 1998; Liebeskind, Oliver, Zucker, & Brewer 1996). However, the networks in life sciences and biotechnological industry are not closely linked or rather, the links are not primarily through easily visible channels such as co-publishing, referencing literature, or co-patenting. There are other channels through which knowledge transfer is effected, including spin-off activity, where scientists are involved in founding new firms, consulting directly to industry, or mentoring of industry projects (Murray, 2002). Strong linkages between the different channels have recently been exposed: in France a significantly higher percentage of life science PhDs embarks on a career in industry from those laboratories which have the strongest working relationships with this sector (Mangematin & Robin, 2003).

Empirical analyses have shown that several of the characteristics of research groups influence their research productivity. We will first discuss variables describing the internal structure of research teams and then further variables such as size, resources and culture which all can be related to the research productivity of a research group.

2.1.2 Composition of research groups

The internal structures of research groups differ, and four different personnel groups have been discussed in regard to their effects on research productivity: (1) highly productive scientists as team leaders, (2) doctoral students and young researchers in general, (3) postdoctoral researchers, (4) other permanent researchers and non-research personnel. Moreover, the diversity of research teams in regard to two further criteria has been found influential: firstly, the variety of disciplines included in a group and secondly, the cultural and national backgrounds of the team members.

Participation of highly productive scientists

One of the group features which affect the overall number of publications of a research team is the participation of a highly productive scientist in such a team (von Tunzelmann et al., 2003). Lotka (1926) established a power law for the distribution of scientific publications on the corresponding authors: Few authors produce a large percentage of the publications whereas the majority of the authors produce only few. Lotka’s law has implications for the publications of research groups: publication figures are higher for groups with such highly productive scientists. The figures might be considered as skewed or not representative for the entire group (von Tunzelmann et al., 2003).

However, in a way the group might also be responsible for the high productivity of the individuals. Highly productive scientists might have more liberties to choose their work places and they might be attracted to similarly productive colleagues which would lead to a cumulative and self-selective appointment process (Bonaccorsi & Daraio, 2003). Long and his co-authors (Allison & Long, 1990; Long, 1978; Long & McGinnis, 1981) have demonstrated that the publication behaviour of mobile scientists adjusts to their destination – for instance, scientists moving from less to more prestigious departments publish more (see also section 2.3).

Moreover, prolific and renowned scholars might attract the best students and postdoctorants which hope to increase their career opportunities through working with an eminent scientist. A recent study on the labs of the University Louis Pasteur in Strasbourg found post-doctoral researchers in particular at the eminent labs (Carayol & Matt, 2004). This finding applied to foreign post-docs even more.

Doctoral students and age effects

In life sciences, as elsewhere, doctoral students and PhDs constitute a sizeable segment of the human resources. For instance, in French life sciences labs, PhD students are estimated to account for nearly 30% of qualified manpower (Mangematin & Robin, 2003). Laredo (1999, 2001) obtains a similar share for labs in human genetics in six different European countries. If technicians are included and the overall lab size is calculated, still one fourth of the human genetics personnel are PhD students. A similar percentage can be calculated from the data on the labs at the University of Louis Pasteur Strasbourg provided by Carayol and Matt (2004): 32% of their scientists are PhD students.

Working with graduate students, post-docs and untenured junior colleagues has been identified as a specific collaboration strategy for established researchers, one distinct from, for instance, cooperating with colleagues with particular complementary skills (Bozeman & Corley, 2004). The value of senior researchers for research groups lies in their research experience, contacts in the scientific community, and access to funding. Doctoral students usually don't have these assets, but they were trained with the most recent knowledge, are not burdened by old opinions and obsolete knowledge, and are particularly motivated as they are still at the beginning of their careers (Levin & Stephan, 1991; Mangematin & Robin, 2003; Stephan & Levin, 1997). Hence, a division of labour between older and younger scientists occurs (Laudel, 2001). Doctoral students contribute their specific and still limited knowledge to the exploration of scientific anomalies (Shinn 1988 cited in Mangematin, 2000) and do the experimental or other empirical work. The team leaders make the theoretical work and help with the interpretation of the evidence (Laudel, 2001).

PhD students raise the visibility of their supervisor through increasing the possible amount of work and publications, and through their subsequent careers (Mangematin, 2000). By obtaining permanent positions in academia, they can become external collaborators of their PhD supervisor, or by moving to industry they can become a channel for knowledge transfer (Dasgupta & David 1994; Mangematin & Robin, 2003). In exchange to these services they receive research training and help of the PhD supervisor to find a new job after the completion of the PhD; this has been described as an implicit contract (Stephan & Levin, 1997) governing the relationship between PhD students and their supervisors. Accordingly, PhDs can be interpreted on the one hand as a resource for scientific and technological undertakings and on the other hand as a "product" of science (European Commission [EC], 2003; Mangematin, 2000).

However, this implicit contract between PhD students and their supervisors tends to be broken, as Stephan and Levin (1997, 2001a) argue for the USA. Above all, the public research sector doesn't offer sufficient permanent positions for the PhD graduates. Where PhD students are considered as cheap labour (caused by their low salaries, funding regulations etc.) the quality of their education will suffer. This is a particular problem of the life sciences and other laboratory sciences, as Hagstrom already illustrated in 1965: 57% of PhD recipients from microbiology, biochemistry, and biophysics – a higher percentage than in any other discipline – agreed with the statement that "major professors often exploit doctoral candidates" (Berelson, cited in Hagstrom, 1965, p. 134). Another breach of the implicit contract occurs because university scholars tend to have insufficient knowledge of private business research, if they don't cooperate with firms (Mangematin & Robin, 2003); hence, they can't teach their PhD students the knowledge they need for being successful in private research.

The significance of doctoral students in research has been analysed either directly or indirectly through looking at the effects of age on group productivity. The latter studies are more common. They usually analyse at the individual level the proposition that science is a young man's game. Levin and Stephan (1991) call it a life-cycle effect on research productivity: senior scientists have fewer incentives to invest their time in research than more junior colleagues, as they will reap fewer benefits before they retire. In other words, as the age of scientists rises the attractiveness of doing research drops. However, the empirical evidence does not unanimously corroborate negative effects of ageing (Barjak, 2005; Fox, 1983; Hemlin & Gustafsson, 1996; Knorr, Mittermeir, Aichholzer, & Waller, 1979a; Levin & Stephan, 1991; Ramsden, 1994; Wray, 2003). Rather peak productivity is reached at mid career between the age of 35 and 45 years; then it levels off gradually. Wray, who investigated at what time in their career eminent scientists made their major discoveries, concluded: "… it is the middle-aged scientists who produce proportionally more than their share of revolutionary discoveries" (Wray, 2003, p. 145). Stephan and Levin (1997) find this curvilinear relationship between age and research productivity in life sciences disciplines.

A related effect of age is discussed by Bonaccorsi and Daraio (2003): they find that a higher average age of all group members has a depressing effect on group productivity. They explain this effect with an underlying lower turnover of personnel, less attractiveness to younger scientists and less scientific vitality. Also, average age of the group members is positively related to the time of existence and size of the group. Interestingly, the relationship between average age and productivity is a lot more pronounced in the basic disciplines geology, chemistry and physics, whereas it tends to disappear in the more applied disciplines of engineering, life sciences and agriculture. Bonaccorsi and Daraio (2003) propose that in the latter disciplines the importance of external orientation and practical experience might prevent the ageing effect.

Direct empirical evidence on whether doctoral students increase the productivity of a research group is scarce. Baird (1986) gives positive correlation coefficients for variables measuring the size of department's doctoral degree programs and their research productivity. However, he states also that "it is unclear whether the presence of able, motivated students promotes research or whether research oriented departments attract and admit such students" (Baird, 1986, p. 222). Cohen (1981) did not find any significant effects on publication figures for the number of "Medical Research Council students" – a category that is not explained further. Carayol and Matt (2004) do not find any significant relationship for the 80 French labs of their study. Laredo and his collaborators (1999, 2001) indicate that in human genetics labs which are highly involved in research (and productive) doctoral students constitute on average 35% of all team personnel which is ten percentage points more than for the overall sample of labs. Pravdic and Oluic-Vukovic (1986) provide an interesting interpretation for scientists which – though they publish a lot and have many co-authors – do not reach their productivity potential: their collaborators are mainly PhD students which produce experimental results but hardly help with the interpretation of the results and the production of the papers.

Postdoctoral researchers

"A faculty member is only as good as his or her best postdoc." With these words cite Stephan and Levin (2001a, p. 682) an associate dean at a medical school of an US American research university. The significance of post-docs is somewhat difficult to measure, whenever post-docs are not distinguished clearly from other non-permanent faculty. Carayol and Matt (2004) find that 19% of the scientists in the labs of Pasteur University of Strasbourg are postdoctorants. Laredo and his collaborators (1999, 2001) asked for post-docs as well as for junior researchers in their survey among human genetics labs. Junior researchers in total accounted for roughly 27% of the scientific staff (22% of overall personnel); of this 27% only one third were denoted as post-docs.

Like doctoral students, postdoctoral researchers have been assumed to be particularly motivated as they also have temporary appointments and are still at the beginning of their careers. And like doctoral students they can hope to get a compensation for the hard work during the postdoctoral phases only when they obtain a permanent position later on. Also from the standpoint of team managers short term appointments provide flexibility in the staffing of groups. Personnel can be adjusted more easily to research topics and needed qualifications. However, as Stephan and Levin (2001a) state, these effects of postdoctorants in research are to some extent specific to the US and determined by its academic system and the funding of scientific research. Under different contractual relations and funding regulations things might be different. They conclude that the significance of postdoctoral researchers is still an open question that has to be answered empirically.

The above cited study of Carayol and Matt (2004) finds a complicated relationship between the productivity of lab members and the presence of post-doctoral researchers: average productivity is not correlated with the ratio of post-docs to permanent researchers, but kurtosis and variance of productivity are. They interpret this as an indication of post-docs being strongly attracted to labs with few highly productive researchers. Postdoctorants "seem to value only fame and excellence when choosing labs" (ibid., p. 1089).

Other researchers and non-researchers

In addition to the heads of research teams, PhD students, and postdoctorants, there are three other groups of team members which might affect the research productivity of a laboratory: other faculty, non-research staff (technicians, administrators), and guests and visitors. Empirical evidence on these groups is limited: Cohen (1981) found positive effects on publication figures within three biomedical institutions for the number of senior staff and attached and guest workers. However, he himself notices that in particular guest scientists and visitors might also be attracted to a lab because of its productivity and visibility in the scientific community. According to Carayol and Matt (2004) full time researchers publish more than university professors with teaching duties.

Investigating the role of non-researchers Carayol and Matt (2004) find that the latter personnel category is more or less evenly distributed across the labs, and that it does not bear any relationship to their institutional recognition. However, the ratio of non-researchers to permanent researchers is positively correlated with the research productivity of the permanents. Accordingly, technicians don't cause fame, but they make valuable contributions to research output. In the typology of Laredo and Mustar (2000), which is also based on French labs, the labs with a lot of technical support staff are those with a public service focus.

Variety of disciplines

The interdisciplinary composition of research teams has been proposed several times as one of the major features of modern research (Gibbons et al., 1994; Ziman, 1994). Ziman (1994) argues that interdisciplinarity is necessary in current science as the scale of research has grown, problems require broad knowledge from different disciplines, and to overcome cognitive lock-ins. A strong argument in favour of interdisciplinarity is made by Hollingsworth and Hollingsworth (2000): analysing two case studies of globally leading biomedical research organisations, they suggest that diverse but scientifically integrated staff that is able to communicate and bridge the disciplinary boundaries has been conducive to generating major biomedical discoveries during the entire 20th century. However, they also warn against the danger of "hyperdiversity" that makes communication ineffective (Hollingsworth, 2002). Advantages of interdisciplinary research are numerous (Nissani, 1997); however, some authors also state possible disadvantages, in particular practical barriers (Younglove-Webb, Gray, Abdalla, & Purvis Thurow, 1999).

Some studies have corroborated the finding that interdisciplinarity is particularly common and important in the life sciences (Sanz-Menendez, Bordons, & Zulueta, 2001; Quin, Lancaster, & Allen, 1997). For instance, around 60% of Spanish life science teams consist of scientists from more than one discipline (Sanz-Menendez et al., 2001).

Cultural and national backgrounds

The benefits of differing cultural and/or national backgrounds in research teams are often not more than a sideline in the literature on research productivity. However, the mobility of scientists between countries, but also between different organisations and sectors inside and outside of academia may lead to a transfer of knowledge and cross-fertilization. Mobile researchers are carriers of external knowledge from other cultures, theoretical schools, organisational contexts etc., transmitting tacit knowledge to their new location (Laudel, 2003). They also can function as gatekeepers into their work environment for their previous co-students and colleagues. Stephan and Levin (2001b) convincingly argue that internationally mobile scientists are highly motivated and among the most able of their contemporaries. This is also one of the assumptions of the brain drain debates in some countries (Davenport, 2004; Laudel, 2003). In particular doctoral students, postdoctorants, and younger and non-tenured scientists are mobile – Mangematin and Robin (2003) state that each year 300 out of 2,000 French life sciences PhDs move from academia to the private sector, whereas tenured scientists hardly leave the academic world.

Hollingsworth and Hollingsworth (2000) stress the contribution of foreign permanent staff as well as visiting scientists to Rockefeller Institute’s scientific success. However, this contribution might have been due to their scientific eminence as well. The investigation of Stephan and Levin (2001b) on the contribution of foreign-born and foreign-educated scientists to US science obtains the result, that both groups of scientists make more exceptional contributions to science than can be expected from their share among the scientific workforce. An older study of Aran and Ben David (1968) showed that post-doctoral research training at foreign universities increased the research productivity of medical researchers. However, this result was probably at least partially due to the particular organisational setting of Israeli medical education which at this time did not provide profound research training and required a re-socialization of the medical researchers abroad.

Mobility does not necessarily lead to a higher research productivity. Long (1978) and Allison and Long (1990) showed that the publication activity of mobile scientists adjusted to the publication activities and prestige of their destination: scientists who moved up to universities with a higher prestige rating published more, and scientists who moved down published less few years after the move.

2.1.3 Further group-related influences on research productivity

Further group-related influences on research productivity are in particular: group size, age and structure, available resources, climate within the group and effectiveness of leadership and management.

Team size

The size of research teams is not as easy to measure as it may seem at first glance. It depends on the definition of a research team that can follow different approaches (see section 2.1.1): an institutional approach taking all researchers in a particular team (Cohen, 1981); an approach based on actual joint research (UNESCO study see Andrews, 1979); or an approach based on co-authorships (e.g. Bordons & Zulueta, 1997; Seglen & Aksnes, 2000). According to these differences of defining research groups, results for group sizes in the life sciences differ:

- Cohen (1981) analyses research groups within three biomedical institutions (two from the US and one from the UK) and obtains mean lab sizes between 10 and 20 members (including senior staff, students, and guests and visitors but excluding technical and administrative staff).
- Bordons and Zulueta (1997) identify Spanish biomedical groups on the basis of co-authorships. Authors were included in a group, if they published more than 60% of their publications with the most productive author of the group. They obtained average group sizes of 8 (pharmacology and pharmacy) and 11 (cardiovascular systems) authors.
- Seglen and Aksnes (2000) employed a similar approach for Norwegian microbiology but included authors already in a group, if they had published a majority of their articles with another more productive member of this group. They get an average group size of 24 authors.

The above listed studies do not corroborate a positive effect of group size on group productivity: Cohen (1981) did not find any relationship between lab size and number of publications. Bordons and Zulueta (1997) even indicate a negative relationship between team size and team productivity in Spanish pharmacology and pharmacy research. Seglen and Aksnes (2000) do not find any correlation between group size and research productivity at all. Hollingsworth and Hollingsworth (2000) argue that growth leads to differentiation and bureaucratisation which are detrimental to making major scientific discoveries even if the organisations are highly productive.

Positive, negative or no significant correlations between group size and productivity have been found in other studies (Arora, David, & Gambardella, 1998; Carayol & Matt, 2004; Johnston, 1994, Knorr et al., 1979a, Stankiewicz 1979; von Tunzelmann et al., 2003). Some point to a curvilinear effect of increasing productivity up to a certain threshold and decreasing productivity beyond this threshold (Johnston, 1994; Stankiewicz, 1979). However, this threshold, i.e. the optimum group size, also varies. Based on his investigation of the combined effects of group size, age, cohesiveness and leadership, Stankiewicz (1979) argued that size only plays a role for specific constellations of the other three variables. In particular, groups larger than seven scientists show a decline in performance, but only if they have inexperienced leaders and low degrees of cohesiveness. For cohesive groups with experienced leaders the effects of size on performance might actually be reversed and positive.

Material resources of the group

The availability of resources is related to the size of a group: the bigger a group is, the more resources it needs. Material resources such as funding, instrumentation and literature receive high ratings in surveys (Baird, 1986; Ramesh Babu & Singh, 1998). However, the UNESCO study listed as one of its general and consistent findings that no relationship between economic and physical resources and the effectiveness of research teams could be found (de Hemptinne & Andrews, 1979). Johnston (1994) presents an interesting hypothesis on this issue that is based on his interviews with Australian researchers: accordingly not the local resources are important for productivity, but rather the access to international centres with state of the art equipment and other researchers which use it there. This entails a shift of focus from local and material resources to the integration into non-local research networks – the main theme of the NetReAct research project. Arora, David and Gambardella (1998) point to a different problem in regard to including funding data in productivity estimations: if funding agencies observe and consider characteristics of the team in the funding decision which also influence the productivity of the grant applicants, then the funding is not an exogeneous factor in the research productivity function anymore and the estimates will be biased.

In regard to the structure of funding it has often been assumed that research funding from private firms might reduce the publicly available research output (and thus harm scientific progress) due to intellectual property concerns and publication restrictions. However, neither the data for labs (Laredo, 1999, 2001) nor for individual researchers (Barjak, unpublished data) can corroborate this. Partially this might be due to the fact that the percentage of funds from private businesses is rather small: only 8% on average in Human Genetics labs (Laredo, 1999) and 8% for chemists, 5.5% for computer scientists, and each less than 2% for astronomers, economists, and psychologists (Barjak, unpublished data) come from private firms.

Group age

Stankiewicz (1979) analyses the effect of group age on performance. His results show a relationship between group age and the cohesiveness of a group and its leadership: group age can have a positive (!) effect on group performance if the group shows low values for the cohesiveness indicators and if it has an inexperienced leader. Groups with high cohesiveness and under the direction of experienced leaders do not show any influence of the age variable. A more recent analysis also included a variable for the lab age that, however, was not clearly related to publication activities (Laredo, 1999).

Group interactions

The positive effects of group-based research stem on the one hand from a mere division of labour and on the other hand from the usage of intellectual synergies and the cross-fertilisation of the team members (Allison & Long, 1990; Stankiewicz, 1979). These effects are determined by the appropriate size and structure of the group but also by the quality of the interactions between the group members. With group leadership and climate the literature discusses two important aspects of group interactions:

- Group leadership: Knorr, Mittermeir, Aichholzer, and Waller (1979b) showed that satisfaction with group leadership and an index on research performance are weakly correlated. Stankiewicz (1979) argued that groups led by young and inexperienced leaders have to overcome several disadvantages such as the leaders' own career interests and fewer external contacts. Fox (1983) defines the term "organizational freedom" as the liberty to select the content and the administrative parameters of the research. Organizational freedom tends to increase publication-productivity, though some studies also stress the beneficial effects of some degree of organizational coordination (Fox, 1983; Bland & Ruffin, 1992). Bland and Ruffin (1992) highlight the fact that the leader of a research group and his or her style of leadership is one of the most important factors for determining research productivity: research-conducive leadership is one that is based on participative governance and decentralised decision-making and at the same time oriented towards clear goals with a clear-cut emphasis on research (see also Ramesh Babu & Singh, 1998). In the same vein, Hollingsworth and Hollingsworth (2000) identify a "leadership which gave particular attention to the creation and maintenance of a nurturing environment, though with rigorous standards of scientific excellence" as particularly important for excellence in biomedical research in their case studies of Rockefeller Institute and Cal Tech.
- Group climate: The UNESCO study operationalised group climate through different items referring to the spirit of innovation, commitment, openness to new ideas, degree of cooperation and communication (Knorr et al., 1979b). It found a positive relationship to research performance.

2.2 Research collaboration and productivity

The previous chapter discussed the influences on team productivity stemming from its structure, size, resources, and culture. However, research groups don't exist in a vacuum but they are integrated into a network of institutions from science and other social systems (see section 2.1.1). In particular, they collaborate with other research groups whenever this entails a benefit for their work. Though research collaborations between research groups have become more widespread in recent decades, and in some disciplines they are the rule rather than an exception, the empirical evidence is still scarce:

"Indeed, the study of research collaborations predominantly or exclusively formed among public sector research entities in the physical, engineering, medical, social and behavioural sciences, remains essentially in its infancy." (David & Keely, 2001, p. 2)

This chapter reviews the most important contributions to the study of research collaborations with a particular focus on their impact on the productivity of research groups.

2.2.1 What is research collaboration?

The term research collaboration is usually used for situations in which individual researchers or groups from academia, public labs or private firms work together to produce new knowledge (see e.g. Katz & Martin, 1997). We limit the further discussion of research collaboration on collaborations in which university scientists participate.

As Katz and Martin (1997) note, the concept of collaboration becomes meaningless if too broad an understanding of joint work is used. They see the duration and intensity of the joint work, the contribution of main elements of the research, and the responsibility for key steps including the initiation and fund raising as main criteria to separate collaborators from non-collaborators. Laudel (2001) distinguishes at the group level five different types of scientific research collaborations based on the content of the collaborators' contributions:

- Collaborations involving a division of labour are understood as collaborations in which all involved groups do creative work.
- In service collaborations one party contributes not creative but only routine work, such as preparing or carrying out the measurements and data collections.
- Provision of access to research equipment is self-explaining; in this case, no further work is requested from the owner of the equipment.
- Transmission of know-how means that procedural knowledge is transmitted, often upon request and in the form of consultancy.
- Mutual stimulation is not related to a specific project or problem, but a side effect of scientific communication.

Only the first of these five collaboration types regularly leads to co-authorship rights for the collaborators (Laudel, 2001). Thus, if co-authorships are used to indicate research collaborations, only a part of all existing collaborations is actually measured. However, these collaborations may be considered as the most intensive and active ones which also include a fair amount of creative work of all the involved parties (Laudel 2001; Luukkonen, Persson, & Sivertsen, 1992).

2.2.2 The impact of research collaborations on research productivity

Research collaborations and research productivity

An enhancement of the cognitive capacities of research teams through knowledge sharing and transfer is one of the possible effects of research collaborations that might increase team productivity (Katz & Martin, 1997). Access to equipment and funding are other possible effects of collaborations which also influence scientific output. Moreover, multi-author papers may be accepted more easily by the gatekeepers to scientific journals, as some of the discussions and counter-arguments already took place during the writing of the paper (Gordon, 1980, cited in Katz & Martin, 1997). In general, the discussions among the co-authors might boost creativity (Katz & Martin, 1997; Laudel, 2001). However, collaboration is not gratuitous, but it creates costs, too, for instance for travel, communication, coordination, administration etc.

Katz and Martin (1997) conclude in their literature survey that the empirical evidence backs the idea of a positive relationship between collaboration and research productivity. Most evidence on the effects of collaboration on research productivity is available at the level of the individual scientist. Pravdic and Oluic-Vukovic (1986) analysed the effects of collaborations among Croatian chemists and found that collaborating scientists produced more publications. Gupta and Karisiddappa (1998) found that average productivity per author in theoretical population genetics is higher in a set of collaborating authors than in a set of non-collaborating authors, employing a full counting method (i.e. co-authored papers are counted with a value of 1 for each co-author). No greater productivity of psychologists with a collaborative research style is found in another, survey-based analysis (Kraut, Egido, & Galegher, 1990). Braun, Glänzel and Schubert (2001) and Glänzel (2002) show, that the relationship between collaboration (average number of co-authors) and publication activity is primarily field dependent: mathematicians with one or more co-authors (on average) do not publish more than those without co-authors; life scientists and chemists with co-authors, however, publish more than the predominantly single authors in their disciplines. The peak publication numbers are reached by scientists with 2-3 co-authors (chemistry) and 4-5 co-authors (life sciences). Scientists with more co-authors than this tended to produce fewer publications. A curvilinear relationship is also found in Barjak (2005): for scientists from seven European countries and five disciplines the relationship between the number of publications and the number of collaborators is first rising and then falling. Evidence at group level is scarcer: Price and Beaver (1965) showed that the size of a collaborative network might have a positive relationship with the members' research productivity: the more co-authors are in a group, the higher was the (fractionally measured) productivity of its members. Arora et al. (1998) also found a positive effect of the number of research collaborations with foreign non-profit institutions on the productivity of Italian molecular biology and genetics research groups.

Newman (2000) discovered that collaboration structures in the life sciences differ from some other disciplines. He studied co-author relationships in biomedical research with a social network approach. His main finding is that rather small networks of co-authors are typical in biomedicine compared to large networks in the other disciplines. However, some of these small networks have joined efforts to contribute to clearly defined, large-scale and long-term goals, such as the Human Genome Project (HGP). Moreover, clustering (meaning the degree to which a scientist's collaborators collaborate with each other) was less widespread than in physics or computer science.

Collaborative research does not only create the hope of achieving higher research productivity, it is also connected to a higher impact of the produced work. This higher impact might be explained through a broader diffusion and higher visibility of the results (Katz & Martin, 1997). The literature reviewed by Katz and Martin (1997) confirms this positive effect of research collaborations on the impact of publications: co-authored papers tend to have a higher impact than single-author papers (e.g. Narin, Stevens, & Whitlow, 1991).

Motivations for taking up a research collaboration

Though the economic perspective on science has repeatedly pictured it as a system based on the competition between scientists for obtaining recognition and rewards (Dasgupta & David, 1994; Hagstrom, 1965), it has not been doubted that several incentives for collaboration exist. These incentives or motives for research collaborations are a good starting point for analysing the impact, as motivations usually stand at the beginning of collaborations and determine lots of their later features. The literature on collaborations lists innumerable different motivations which have cognitive, social, financial or teaching purposes underneath. Among the most important motives are probably (see e.g. Beaver, 2001; Bozeman & Corley, 2004; Katz & Martin, 1997; Katz, 1994; Kraut, Egido, & Galegher 1990):

- *Cognitive purposes:* Basically the access to other scientists' expertise and knowledge, for instance in order to pool resources for solving large and complex problems, to obtain specialised knowledge, to realise a division of labour and facilitate a further specialisation, to learn a new method or the use of new instruments, etc.
- *Social purposes* are often the increase of prestige and visibility through working with renowned colleagues and the fun and pleasure of collaborative work increasing work satisfaction. Also, the maintenance of pre-established relationships might be another social purpose.
- *Financial purposes:* As research has become more and more expensive and resources have not grown to the same extent, the access to equipment, resources, data and other evidence can be facilitated through collaborations. Moreover, funding regulations might require collaborative proposals.
- *Non-research purposes:* Last but not least, research collaborations may also be taken up for non-research purposes, above all the education of students or junior researchers has been mentioned as such a purpose.

In particular the first group of cognitive purposes should have positive effects on research productivity as they are closest to the actual production of new knowledge. However, goals might not be met, motives might change as research is being done, and the other listed purposes not necessarily have a positive relationship with research productivity. Therefore it is necessary to take a closer look at the empirical evidence on how research collaborations correlate with research productivity.

Structural characteristics of collaborations

Not only the overall amount of collaboration and the motivations that lead to collaborative research influence the work of a scientific research team, but also the characteristics of its collaboration partners. They determine what kind of benefits the collaboration generates. The characteristics that receive the most attention in the scientific discussion of research collaborations are: (1) local versus global collaborations, (2) disciplinary versus interdisciplinary collaborations, and (3) public versus private collaborations.

Ad (1) Local versus global collaborations: Different reports show that the importance of scientific collaboration at international level has grown in the last 25 years (EC, 2003; Narin et al., 1991; National Science Board [NSB], 2002, 2004). However, the level of collaboration varies by scientific specialty and country. Within the life sciences notable differences exist. For instance, US American zoologists have been found to have many collaboration partners and collaborate a lot at international level, whereas the collaboration networks of general biologists and biochemists overall and in particular at international level are rather small (Bozeman & Corley, 2004).

In principle, the benefits listed for international research collaborations are of the same nature as those listed for collaboration in general (see e.g. the benefits listed in Georghiou, 1998); but international collaborations clearly create additional costs, for instance due to the necessity of bridging language differences, differing research cultures, and finding valid contractual arrangements. However, obviously there have to be additional benefits of international collaborations, too, that outweigh higher transaction costs; otherwise it would be hard to explain their general growth within the last decades.

One possible gain of international versus local collaboration is that the matching between needed and actually found resources (knowledge, equipment, objects of study) produces better results. In particular, in small countries the available range of research capacities is not as broad as in large countries, and certain skills or equipment might not be available at all (Luukkonen et al., 1992; Thorsteinsdottir, 2000). These gaps can be closed through finding a foreign partner. But teams from small countries might themselves be attractive partners, too, for instance if they have unique local resources or objects of study to offer (Georghiou, 1998; Thorsteinsdottir, 2000). Large scale problems of global relevance induce national governments and administrations to support and provide the resources for international research collaborations. Moreover, scientific specialisations vary across countries and collaborating with a partner from a stronger country might lead to more intensive knowledge transfer and learning (Georghiou, 1998). This means that on average teams in relatively weaker countries profit disproportionately more from collaborations with relatively stronger countries.

The empirical evidence on the consequences of international research collaborations for research productivity is mainly supportive. The study of Bordons et al. (1996) on Spanish biomedical research shows that international collaboration raised the productivity of the team leaders and the impact of the published work. According to Narin et al. (1991) their finding that biomedical papers with international co-authors have a higher impact than single author papers and nationally co-authored papers can be generalised to other disciplines. This is shown in Glänzel (2001) for a range of countries. Glänzel & Schubert (2001) found that co-publications of several country pairs may attract fewer citations than expected on basis of the corresponding domestic reference standards, and called this type of co-publication links cool links. Unlike in biomedical research, where the observed citation impact of most analysed countries was higher than the domestic impact of at least one of the involved partners, and often higher than the world standard, too, the attractiveness of joint papers of some countries was unambiguously low in chemistry and mathematics. International co-authorship seems, therefore, not always to pay for all partners. Persson et al. (2004) have shown that international collaboration has in general a more pronounced positive effect on citation impact than local or domestic collaboration has. Contrary to their own expectations, Carayol and Matt (2004) find that the share of internationally co-authored publications and research productivity are not correlated.

Ad (2) Disciplinary versus interdisciplinary collaborations: The growing importance of interdisciplinary research has already been reported above (see page 22). Analysing Spanish research teams in two biomedical fields and in material science, Sanz-Menendez et al. (2001) have shown that collaboration between teams from different disciplines is more common if the research teams consist of scientists from a single discipline. Hence, interdisciplinary collaborations within and between research teams are in a way substitutes. We do not know of any analysis that tried to measure the effects on interdisciplinary research collaboration between teams on research productivity.

Ad (3) Public versus private collaborations: The literature on collaborations between public research organisations and private firms is by far too broad to be reviewed here in total. The topic has been discussed for instance in the economics of innovation literature (Cohen, Nelson, & Walsh, 2002; Gibbons & Johnston, 1974; Kline & Rosenberg, 1986; Mansfield, 1991, 1998; Salter & Martin, 2001); in the interdisciplinary literature on technology transfer (Bozeman, 2000; Charles & Howells, 1992; Lee, 1996; Rahm, 1994; Rogers, 2002; Williams & Gibson, 1990); and in different strands of sociological and interdisciplinary science studies (Etzkowitz & Leydesdorff, 1997, 2000; Gibbons, et al., 1994). And this list is certainly only the tip of the iceberg. Often, this literature focuses on the social and economic consequences of scientific research. Also, it discusses a wealth of channels or carriers through which the exchange of knowledge is realised: informal contacts, joint and contract research, scientific publications, conferences and the hiring of scientifically trained graduates have been identified as particularly important (c.f. Arundel, van de Paal, & Soete 1995; Meyer-Krahmer & Schmoch, 1998; Fritsch & Schwirten, 1999; Cohen, et al., 2002). However, we deal here with a narrower problem and in a way from the opposite perspective: What are the effects of research collaborations between science and industry on the research productivity of the scientific partners, in particular research teams?

As Brooks (1994) and others have pointed out, industry benefits from science, but science also benefits from industry. For instance, scientists may obtain new research questions or instrumentation and techniques from the exchange with private firms.

Private firms provide funding for scientific research projects and open their doors for students at all levels to implement their knowledge and practice their skills. However, concerns have also been voiced: because of private funding public research may stray away from the more basic type of research problem to the more applied type; firms might introduce restrictions to the disclosure and publication of research results, because they want to appropriate them first. This produced the fear of negative long term effects on the speed of scientific progress. Do the benefits or the costs of public-private research collaboration and knowledge exchange prevail?

The evidence rather points to the dominance of the benefits. Pavitt (1998) looks at the issue from a macroeconomic point of view. He argues that the growth paths of the East Asian "Tiger Countries" and, even more clearly, the developments of Germany, Japan, and the UK since the Second World War strongly suggest that technological progress precedes scientific progress. Other authors look at evidence at the micro level: Hicks and Hamilton (1999) show that papers with co-authors from science and industry are more often cited than single-university papers; so, the impact of this inter-organisational research is higher than the impact of research carried out within a single university. However, they also state that the private agenda seems to dominate the character of research which is indeed of a more applied type. A positive effect of the collaboration between academic research and private business research on research productivity is found for French labs (Carayol & Matt, 2004). Joly and Mangematin (1996) point to the negative aspects of public-private research partnerships. In particular, they state that a dependence on private funds may interrupt and burden research processes; moreover, they argue that not all types of science-industry research relationships may trigger learning processes for the scientific partner. This depends essentially on the logic of the relationship, i.e. whether it is governed by proximity and trust, market mechanisms, or pre-competitive clubs.

All in all, the empirical evidence seems to support a positive influence of research collaborations, and in particular international collaborations and collaborations with private firms on research productivity. In addition to collaborations and team characteristics the literature on research productivity points to further factors which will be discussed in the next section. These factors are the characteristics of the parent organisation, the scientific discipline, and the research system of a country.

2.3 Further environmental covariates of research productivity

Characteristics of the parent organisations

The parent organisations of research teams determine the teams' activities in many different ways.[1] They may either refer to formal regulations, settings, or missions, or to informal circumstances such as the climate and collegiality.

- Formal issues: The capacity to make decisions in regard to strategic issues such as funding, personnel, and research topics certainly constitutes an important influence on the performance of research teams. The recruitment of professors and other senior staff, is often done at other hierarchical levels, whereas team leaders at least participate in the decision making process when it comes to the recruitment of junior researchers and doctoral students (Laredo, 1999). The general mission or focus of the organisation is another factor that affects research. It has been shown in several studies that research productivity is higher in organizations which focus on research (see Baird, 1986; Bland & Ruffin, 1992). In particular the productivity rates are lower in organizations oriented primarily towards undergraduate education (Blackburn et al., 1978). Other issues which affect and regulate the work of research teams are the existence of periodic evaluations, compulsory procedures (like costing rules or authorisations to bid) and support services (e.g. patent office, EU contact office) (Laredo, 1999).
- Informal issues: Different authors (Allison & Long, 1990; Fox, 1983; Hollingsworth & Hollingsworth, 2000; Long, 1978; Ramsden; 1994) claim that a research-encouraging environment, exchange and communication, and general collegiality within an institution enhance productivity. Ramsden (1994) shows that the extents of departmental and individual research activities are correlated; moreover, a cooperative environment plays a crucial role. In several different publications Long and his co-authors have revealed that the effect of departmental affiliation on research productivity is stronger than the effect of productivity on departmental affiliation (Allison & Long, 1987, 1990; Long, 1978; Long & McGinnis, 1981). They showed that the publication activity of scientists (including biochemists and biologists), who had changed their affiliation, adjusted to the prestige of their new department after few years. Scientists moving to more prestigious departments became more prolific, whereas scientists moving to less prestigious departments published less. They propose different hypotheses to explain this phenomenon: superior departments might have the prescience to select promising researchers; other departmental characteristics than prestige like the availability of resources and facilities, intellectual stimulation and motivation might cause the positive effect on publication activity; the higher publication figures might be due to a bias in the

[1] We write „organisations", because there might be several hierarchical levels which more or less exert an influence, such as the university, school, faculty, department, institute, or laboratory.

publication system that helps scientists from prestigious departments. In particular, they state:
"These results suggest that facilities affect the number of publications while other factors (motivation and intellectual stimulation) have greater impact on the "quality" (as measured by citations) of the work published." (Allison & Long, 1990, p. 477)

Scientific disciplines and fields

Whitley (2000, p. 81) speaks of scientific fields as "a particular kind of work organization which structure and control the production of intellectual novelty through competition for reputations from national and international audience to contributions to collective goals". He advances two main dimensions which cause differences between scientific fields in regard to the production of new knowledge: (1) the degree of mutual dependence between researchers that is higher if the work of one group depends very much on what is done in other groups; (2) the degree of task uncertainty in producing and evaluating knowledge claims refers to how established work procedures, problem definitions, and theoretical goals in a field are. Besides these differences of work organisation, there are further differences between disciplines and fields such as the levels of interpersonal recognition amongst community members, size of field, and institutional frameworks (Fry, 2004). These also influence where and how new knowledge is publicised, in journal articles, books, at scientific conferences, on internet databases etc. The effects are particularly powerful, when social sciences and humanities are compared with the natural sciences (Hicks, 2004). However, the new communication networks also change the communication models within fields of the natural sciences in different directions (Hilgartner, 1995; Kling & McKim, 2000).

According to Prpic (1996) there have been relatively few cross-disciplinary studies of scientific productivity. She herself differentiates between natural sciences, technical sciences, biomedical sciences and social sciences and humanities and finds that the number of scientific publications is dependent on the discipline. In another analysis of chemistry, history, and psychology departments, Baird (1986) also finds differing covariates. Barjak (2005) and Barjak and Harabi (2004) compare astronomers, chemists, computer scientists, economists and psychologists and find significant differences, too.

Country

That per capita rates of scientific publications vary significantly between countries is a well known fact: For instance, according to the latest European Report on Science and Technology Indicators, from 1996-99 a researcher in Switzerland published on average 2.24 scientific articles, in the UK 1.65, in Germany 0.99, in the US 0.86 and in Japan 0.46 (EC, 2003). However, why this is so and in particular why productivity differences between countries with well developed research systems are of such a magnitude – for instance between Switzerland and Japan the relation is 5:1 – is still a more or less unresolved puzzle. Differing specialisations of the national research and innovation systems (EC, 2003) and a bias of the publication data towards the English language (Van Leeuwen, Moed, Tijssen, Visser & van Raan, 2001) are some tentative explanations.

Felderer and Obersteiner (1999) find diseconomies of scale at country level, i.e. additional resources produce more research output, but the increase per resource unit is declining. They explain this disadvantage for larger countries with less openness and a greater focus on the domestic market, complex and less flexible science systems, and congestion and invention exhaustion. Narin et al. (1991) furthermore point to a rather close correlation between GDP and scientific output. Numerous other influences are possible. For example, the systems of research funding and evaluation differ in regard to the emphasis they put on international publications as an output of this funding. The British Research Assessment Exercise is based on an elaborate evaluation scheme that includes an emphasis on publications that is not present in most other European countries. The funding system and labour market regulations also act on the structure of research groups and their productivity (Stephan & Levin, 2001a).

Empirical work on research teams has often been limited to individual countries, because of the differences of finding internationally comparable units. A recent EU-funded project, however, undertook the effort (Laredo, 1999, 2001). One of its notable findings is that activity profiles – that is the importance of research, training, industrial, and clinical involvement – of human genetics labs hardly differ between countries. The author concludes:

“Does this mean that we should, at least partly, question assumptions, which focus on national differences? The test on labs in Human Genetics, a very specific field indeed, would suggest so, since, once created (including the allocation of research staff), there do not seem to be major differences between countries in favouring one or another of the profiles.” (Laredo, 2001, p. 292)

2.4 Summary of the literature review, hypotheses and operationalisation of the hypotheses

The present section summarises the results of the survey on the factors that determine research productivity. It formulates hypotheses and includes a preliminary operationalisation within the NetReAct project. Both, hypotheses and operationalisation will be revised as the work proceeds.

2.4.1 Measuring research productivity

Science produces a range of different outputs: skilled graduates, new instruments, new methods, prototypes and publications. However, it is not possible to measure all products with equal ease. For instance, skilled graduates are a co-product of scientific research and higher education and it is difficult to separate the contributions of each activity, because they are often delivered by the same person, the university professor, and sometimes even on the same occasion. Assessment of research productivity in the NetReAct project will be based on two approaches: bibliometric data and webometric data. Their strengths and weaknesses will be discussed in more detail in the deliverables of work-package 2. The following section only attempts to give a general guideline.

Bibliometric data

Scientific publications are one, if not the, most important outputs of scientific research. They partially capture the essence of other output forms and contain the theoretical knowledge that constitutes the base for many discoveries (OECD, 2001). Data on scientific publications and citations are only collected on a regular basis by the US-based firm Thomson ISI (formerly Institute for Scientific Information [ISI]). This database has been found to contain some weaknesses in regard to representativeness (EC, Research DG, 2002; NSB, 2002; OECD, 2001; Leeuwen, et al., 2001). The mere number of publications doesn't say anything on their quality and the actual contribution of different co-authors (Knorr et al., 1979a). Although objections have been raised against the journal coverage and the data processing policy of the ISI in preparing their database, its unique features meet basic requirements of bibliometric technology. Among these features we mention:

- Multidisciplinarity: All research fields in the life sciences, hard sciences, mathematics and engineering are represented.
- Selectiveness: Periodicals covered by SCI are chosen on the basis of quantitative criteria (impact measurements), and the selection is generally reinforced by expert opinion.
- Full coverage: All papers published in periodicals covered by the SCI are recorded.
- Completeness of addresses: The addresses of all authors are indicated, allowing analyses of scientific collaboration and the application of full publication counting schemes.
- Bibliographical references: References are processed with each document. Redefining references as sources makes it possible to analyse citation patterns and to construct citation indicators.

Citation data gives an insight into the quality of the scientific output. The number of citations which a paper receives reflects its influence on the development of the field and the stock of knowledge (EC, 2001; NSB, 2002). But citations are also created for other purposes than the quality of the cited paper which limits their value for evaluation purposes (Borgman & Furner, 2002; Cronin, 1987; van Raan, 2003).

The economic concept of productivity relates output to one or several inputs like labour, capital or all production factors in order to measure the relationship between efforts and outcomes of economic activities. Research productivity is usually interpreted as the labour productivity of researchers, i.e. the output per research worker is compared for a certain time period, assuming that the researcher(s) more or less used the same amount of time for research within this period. Within the NetReAct project we look at research productivity basically at institutional and country levels.

Some authors advance the hypothesis that the large scale collaborations like the Human Genome Project have the capacity to transform the role of research networks in the life sciences and how scientific results are communicated (Hilgartner, 1995; Weller, 1996). The more research results are published in data bases or on electronic networks, the lower the validity of bibliometric data to assess research output. These far reaching consequences still have to be investigated in more detail to find out whether they are a manifest tendency or nothing but a “cyberplatonic dream” (Glasner, 1996). Still, in order to take this critique somewhat into account, within the NetReAct project the presence of research groups on the World Wide Web is also investigated and linked to bibliometric measures of research productivity.

Webometric data

NetReAct relies on the emerging science of Webometrics for key input both to sampling and to analytical work, including contributions on assessing the relationship between collaboration and research impact. Webometrics is defined to be “the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the WWW drawing on bibliometric and informetric approaches” (Björneborn & Ingwersen, 2004, p. 1217). This definition emphasises a heritage in the information science field of informetrics, in addition to the scope of studies (the Web) and their methodological orientation (quantitative).

There has been an explosion of investigations into the use of the web for publishing and communicating academic research (Thelwall, Vaughan, & Björneborn, 2005). The new importance of the web for modern research means that no investigation into science communication and networking can be complete without the inclusion of data about the use of the web. Although journal articles and conference papers are still important for research communication, web sites are also used for this purpose, and personal pages can make available resources, such as academics’ CVs, that were previously difficult to obtain in large numbers.

Valuable approaches include comparing web use between and within disciplines, and comparing web use across countries (Heimeriks, Hörlesberger, & van den Besselaar, 2003; Tang & Thelwall, 2003; Thelwall, Tang, & Price, 2003). These investigations can be useful to reveal cases of excellence for others to imitate, and cases where remedial action is urgently necessary.

Web publication based indicators: Web publication counts indicate the number of web pages produced by a given department, counting all types of web page, irrespective of content. Web publication counts are used as a standard measure of web output activity. In some cases, it is more appropriate to count the number of web directories of content, rather than individual pages, but this depends upon the publishing profile of the departments or universities studied (Thelwall, 2002). Note that web publication counts are not proxies for research output, although they correlate with it (Thelwall, 2002b); they encompass a wide range of different activities, predominantly relating to research and education.

Inlink based indicators: The number of hyperlinks pointing to a department is partly an indicator of how valued the content of its web site is, and partly an indicator of how much it collaborates outside of the university (since collaborations frequently result in links being created, for example in project web pages). Inlink counts, in general, are indicators of a range of types of informal scholarly communication. Their most valuable use is as benchmarks, to be used to identify departments and institutions that are not using the web effectively, falling significantly below the benchmark figures.

Bibliometric indicators will be used as the main metrics of research productivity. The value and significance of webometric indicators as productivity measures will be assessed within the course of the project – it can be considered as one of the innovative methodological contributions of the NetReAct project to measuring science. For the sake of brevity, we will limit the hypotheses formulated in the following sections on research productivity in general. As the project develops and the empirical work advances more nuanced formulations will replace the present ones.

2.4.2 Research teams and their influence on research productivity

Definition of research teams in the present study

The scientific literature on research teams usually uses one of three approaches for defining research teams: (1) the institutional approach stresses that researchers of a research team come from the same organisation; (2) the functional approach values the joint work of researchers; (3) the approach based on co-authorship takes joint publications as the criterion for adding researchers to a team.

For the present study, a definition based on joint research was considered the most appropriate one. As one of our main purposes is the evaluation of the role of doctoral students and post-docs in research groups, we cannot limit the groups to co-authors. In particular doctoral students participate in research teams and carry out the labour-intensive groundwork, but they might not publish before finishing their doctoral thesis. The institutional definition would force us to ignore the open organisational approach to research groups taken in some countries. For instance, in France scientists with different organisational affiliations, usually universities and non-university research organisations, join efforts in so called mixed research teams. Hence, we understand as a research team or research group a group of people, scientists and non-scientists, which work in the same location for a longer time period to produce new scientific knowledge. The group of people is part of a larger organisation (university, department, school etc.) and at least some of its members are employed by a university. Also, the team is recognized from the outside as a separate entity.

As the definition shows, we employ different restrictions in regard to location, time, and people. The geographical restriction may be somewhat counterintuitive and against the growing tendency to virtual collaboration facilitated by the internet and other computer networks. The delimitation is nevertheless necessary in order to clearly meet one of the main objectives of the study which lies in juxtaposing the effects of group features and inter-group collaboration on research productivity. The inclusion of non-collocated and virtual members in the groups would probably make it difficult if not impossible to distinguish clearly between external group members and collaborators. For the same reason, visiting scientists and guest workers are not included in the research teams.

Also, this group of people might be attracted to a lab by its good research performance and, in this case, they are a consequence of high productivity but not its source (Cohen, 1981). To exclude short term guests and visitors we set a minimum duration of stay with the group of twelve months. A problem is the categorisation of graduate students at master level. Having in mind that their main purpose is learning and broadening their state of knowledge but not the production of new knowledge, we tend to exclude them. On the other hand, in some groups they might render valuable research assistance and it would be false to ignore this input. We don't know of any literature that has dealt with the importance of this type of assistance in research, and we propose to resolve the problem empirically: how master students will be handled will be decided on the basis of the empirical results.

Another issue should be pointed out: Research teams as we understand them might but need not coincide with what Joly and Mangematin (1996), Laredo and others (1999, 2001) defined as laboratories. Bigger labs which consist of several units or teams are not included entirely, but only with one or several of their teams. This selection of the smaller unit is mainly due to our topic and empirical approach: as we request detailed information on doctoral students, post-docs and collaboration activities, we have to address informants with sufficient knowledge on these issues. Furthermore, Joly and Mangematin (1996) point out that the logic which governs the relationships between science and industry varies between research teams:

"Laboratories are often composed of one or two research teams which have differing types of relationship with their environment. Industrial relationship logics therefore vary within the laboratories." (Joly & Mangematin, 1996, p. 916)

This should also apply to the relationships between research teams from science – it would be daring to assume that all teams within a lab have the same collaboration relationships and degree of integration into scientific networks. Therefore, the team level is essentially the more appropriate one for our type of analysis.

Participation of highly productive scientists

Highly productive scientists can bias the productivity figures of a group. Due to their research productivity these scientific "stars" should have reached the level of principal investigators and heads of research teams. Because of their productivity and reputation they can choose from the best students and postdoctorants and they are sought by many as collaborators. We expect that this affects the structure of the group, the range of collaboration partners, and the overall performance of the group.

> H1.1: The structure of the research teams led by highly productive scientists is more diverse than the structure of teams led by less productive scientists.
>
> H1.2: The collaboration partners of research teams led by highly productive scientists are more diverse than the collaboration partners of teams led by less productive scientists.

H1.3: The research productivity of teams led by highly productive scientists is higher than the research productivity of teams led by less productive scientists.

Operationalisation: The identification of highly productive scientists in a research team can be based on the one hand on the bibliometric data. On the other hand, scientific eminence is also reflected through awards and honorary positions in academia. Barjak (2005) finds that recognition assessed through the responses to a four-item question (asking about awards, service on professional committees, editorial boards, and advisory committees within the previous five years) correlates significantly with research productivity. Hence, the survey will include a similar question to assess the level of recognition of the research group leader.

Doctoral students and age effects

PhD students carry out specific tasks in research which are often centred on the empirical groundwork, whereas the principal investigators and more senior researchers do the theoretical work and interpretation of the results (and the overall research management). Hence, we can assume that the realisation of this division of labour is conducive to research productivity.

H2.1: Research teams that realise a division of labour between junior and senior scientists have higher research productivity than teams that don't realise a division of labour.

Empirical investigations of the effects of doctoral students and youth on research productivity don't permit the conclusion that a general positive effect prevails. Firstly, doctoral students are still research apprentices which have not yet reached maximum productivity. Secondly, the situations of doctoral students differ: on the one hand they might be considered as valuable group members who contribute to the publications of their research group and visibility of their supervisors; on the other hand they might be nothing else but cheap labour. In that case, they replace technicians who are without doubt necessary and important but not a decisive factor to labs' performances. We therefore do not expect a general positive impact of doctoral students on the productivity of research teams.

H2.2: Research teams with many PhD students do not have higher research productivity than teams with few PhD students.

H2.3: The research productivity of teams fulfilling the implicit contract with PhD students is higher than the research productivity of teams that break it.

Operationalisation: In order to assess whether an internal division of labour between PhD students and senior scientists is realised in research teams, a detailed description of the tasks of different personnel groups would be ideal; however, this is beyond the possibilities of the NetReAct project. A possible alternative is an assessment of the age structure of the research team which rests on the assumption that younger members are students and older members are senior scientists. The fulfilment of the implicit contract with PhD students can be assessed in two different ways: (1) through letting the team leaders describe the functions and work situations of their PhD students; (2) through gathering information on the career progression of former PhD students. If PhD students above all work on their thesis and advance the career ladder after finishing the PhD, we can assume that the implicit contract has been fulfilled.

Postdoctoral researchers

The work of postdoctoral researchers in science is governed by a similar implicit contract as the work of PhD students. However, postdoctorants are more advanced apprentices than doctoral students and their tasks and contributions to research differ. Also, according to the life cycle model postdoctoral researchers are closer to their maximum productivity level as they tend to be older than doctoral students. Empirical research is scarce but it points to a positive relationship between the number of post-docs and research productivity. This guides our expectations:

> H3.1: Research teams with many postdoctoral researchers have higher research productivity than teams with few postdoctoral researchers.

> H3.2: The research productivity of teams fulfilling the implicit contract with postdoctoral researchers is higher than the research productivity of teams that break it.

Operationalisation: The fulfilment of the implicit contract with postdoctorants is assessed in the same way as for doctorants.

Other researchers and non-researchers

The role of other faculty, technicians and administrators and guest scientists is not among the main concerns of the NetReAct project. Their numbers will be assessed and used as covariates in the empirical analyses, but no separate hypotheses are formulated and investigated.

Variety of disciplines

We expect positive effects on research productivity for interdisciplinary research teams; we expect this above all for post-doctoral researchers as they should have absorbed the cognitive und cultural content of their original discipline to larger extent than PhD students. Also, due to their greater experience they have been exposed to interdisciplinary research for a longer time and hence they have greater facility in communicating across disciplines.

> H4.1: Research teams with PhD students from other disciplines are more productive than research teams without such students.

> H4.2: Research teams with postdoctoral researchers from other disciplines are more productive than research teams without such postdoctoral researchers.

Operationalisation: Measuring interdisciplinarity of research teams requires an assessment of the disciplines of the team members. This is not as easy as it might seem to be: researchers can have a degree in one discipline but do research in another; or they can have degrees from several different disciplines. As we assume that the research team determines the research disciplines and fields in which the research is carried out, it would probably not produce the desired result to collect information on the fields of research of the team members. Hence, as a first attempt we employ an education-oriented approach and assess the scientific discipline of the team and selected members in the following way:

- research team: disciplines of the PhD programmes to which it contributes according to ISCED 1997 (if it doesn't contribute to doctoral education, graduate education is chosen instead, see annex A-1 on the disciplines)
- head of the research team: discipline of the PhD according to ISCED 1997
- post-doctoral researchers: discipline of the PhD according to ISCED 1997
- doctoral students: discipline of the master degree (or bachelor) according to ISCED 1997

Cultural and national backgrounds

Mobile scientists have been assumed to make positive contributions to research productivity at their target locations mainly for three reasons: (1) they outperform their contemporaries, (2) they are committed to success at their new locations, and (3) they transmit (tacit) knowledge that benefits their new co-workers. Even though we lack an empirical corroboration of the positive effects of mobile team members on the research productivity of their teams, we expect to find such an effect. We investigate it for both, the country of birth and the country where the highest degree (master or PhD) was obtained. As with the disciplinary composition, we anticipate more pronounced effects for postdoctorants than for doctoral students.

> H5.1: Research teams with PhD students from other countries of birth are more productive than research teams without such students.

> H5.2: Research teams with PhD students who obtained their graduate degree in another country are more productive than research teams without such students.

> H5.3: Research teams with postdoctoral researchers from other countries of birth are more productive than research teams without such postdoctoral researchers.

> H5.4: Research teams with postdoctoral researchers who obtained their PhD and/or graduate degree in another country are more productive than research teams without such postdoctoral researchers.

Operationalisation: The countries of birth and of education will be assessed in the questionnaire through separate questions. In line with the tender specifications we use the following country groups:

- the country where the team is located,
- other EU member states (membership as of 2003), Norway, Switzerland
- a new member state as of 1. Mai 2004,
- other Eastern European countries (including Russia),

- the USA
- Canada, Australia or New Zealand
- Japan, China, India, Korea
- any other country worldwide

Further team-related influences on research productivity

We discussed above several control variables for the further team-related influences on research productivity, namely size, material resources, age, leadership, and climate. No hypotheses for these control variables are formulated as they are not the main focus of the study.

Operationalisation: Though the inclusion of these variables seems at least in part necessary in order to obtain valid results for the main variables of the project, a limiting factor to their inclusion is the overall length of the questionnaire and the ensuing consequences for the response rate. Hence, only minimum space can be devoted to these control variables in the survey.

- *Team size* will be obtained through the addition of all personnel working in the team. As it has been shown that “heads” rather than “full time equivalents” are important, we will not make any adjustments for workloads below a hundred percent.
- The evidence on the role of *material resources* for group productivity is rather inconclusive. Also, the measurement is highly complex, due to many different sources of funding, different time horizons of grants and contracts, and last but not least, the underdeveloped financial control mechanisms in many research organisations. Hence, we will include only few control questions on the research team’s sources of funding and the funding of PhD students and post-docs.
- The relevance of *group age* for group productivity is an open issue, too. Nevertheless, a question will be included in the survey, as it is easy to ask and easy to answer.
- *Group leadership*: Critical aspects of group leadership are the age and experience of the team head (Stankiewicz, 1979) and his or her style of leadership (Bland & Ruffin, 1992). The former will be assessed through direct questions. For the latter questions on the forms of governance and decision-making (participative/ decentralised versus centralised) and the setting of goals in the research team will be included in the questionnaire.
- *Group climate*: The UNESCO survey assessed group climate through questions on several issues related to climate. In order to avoid biased answers, the responses of several group members were obtained. This is not possible in the NetReAct survey, and we fear that asking the leaders about the climate in their teams will not produce valid results. Therefore, this issue has to be dropped from the survey.

2.4.3 Research collaboration and research productivity

Research collaboration between teams

Several distinct forms of research collaboration exist and only some of them actually lead to the co-authorship of scientific publications. However, collaborations that lead to co-authorship are particularly intense and important as the involved parties contribute creative work. Therefore, within the NetReAct project we will measure collaboration between research teams by means of co-authorship data. We assume that all publications from the ISI database with co-authors from other organisations than the one in the sample indicate that collaborative research took place. Intra-team collaboration, defined as papers with multiple authors from the research team in the NetReAct sample, will not be considered separately.

Based on this operationalisation of research collaborations we will be able to calculate the shares of collaborative publications of all publications. Though this should produce a valid reflection of the amount of research collaboration of the team, there are some weaknesses to this operationalisation that have to be mentioned. First, if researchers joined a new team before the work was published they might appear with the address of their new team. Also, researchers with two or more affiliations might appear as external collaborators though in fact they are not. In both cases we overestimate research collaboration between teams. Second, scientists may collaborate but publish separately. Here, we underestimate collaborations. Third, we assume that joint creative work of scientists from different teams led to the co-authored publication, but we actually can’t ascertain this assumption. It is well known that co-authorships are given for other purposes than joint work (see e.g. Laudel, 2001). Moreover, we don’t know how much each collaborator contributed and how the quality of the contribution was. Fourth, many types of collaboration that also entail the transfer of knowledge but for one reason or the other do not lead to co-authorship are not recognised through the indicator.

Productivity effects of research collaborations

Economic theories of science usually describe it as a system that is ruled by competition, above all the competition between individual scientists for recognition. Therefore, collaborative research has to bring about additional benefits to the involved parties; otherwise it would not be undertaken. Previous research has mostly investigated the relationship between research collaboration and research productivity at the individual level. In general, it has found a positive relationship stressing the multiple benefits of collaborations. However, some evidence exists that collaboration also produces costs for travel, communication, coordination, administration etc. Empirical results suggests that the benefits and costs of collaboration increase with the amount of collaborative activities: first benefits increase more, but beyond a certain amount of collaboration the costs increase more. All in all, this lets us expect a curvilinear relationship.

> H6.1: Collaborating research teams have higher research productivity than non-collaborating teams.
>
> H6.2: The marginal effect of additional research collaboration on research productivity is decreasing (curvilinear relationship).

Operationalisation: Research collaborations are assessed through counting publications from the ISI publication database with co-authors from other research teams than the team in the sample.

Motivations for taking up a research collaboration

There are probably as many motives to collaborate as there are research collaborations. However, previous research on collaborations has identified motivations which can be attributed to four basic types: cognitive, social, and financial motivations as well as motivations related to other activities in particularly teaching. We do not know of any previous research that has investigated whether the motivations of becoming involved in research collaborations affect the outcomes, but we expect this to be so. In particular, we assume that research collaborations for cognitive purposes bear a stronger positive effect on research output than the other motivations.

> H7.1: If research teams predominantly collaborate for cognitive purposes, then their collaborations will have a positive effect on their research productivity.

Operationalisation: The research team's motivations to become engaged in research collaborations will be assessed through the survey. Different variables will be used to record individual motivations. Through further processing of the variables (e.g. a factor analysis) it should be possible to identify the underlying factors. Of course, research collaborations can be motivated by more than one type of motivation, and a further grouping of the research teams according to the dominant motivation(s) will be necessary in the analysis.

Structural characteristics of collaborations

Local versus global collaborations: International collaboration has grown notably in the past 25 years. Though the empirical evidence is not convincing, we expect that its benefits for research productivity are higher than its costs, as otherwise this growth could hardly be explained.

> H8.1: Research teams which collaborate internationally have higher research productivity than teams which do not collaborate internationally.

Collaboration with "strong" partners means that more knowledge is transferred to the weaker team. According to the relative prominence of the literature (ratio of citations to publications) in 2001, we find that in particular Switzerland, the US and the UK are very strong, Germany, Sweden, France and Norway are of high to medium strength, and Italy, Spain, Portugal, Hungary, and Czech Republic are less strong (compared among the sample countries!) in life sciences research (NSB, 2004). Therefore, we expect that on average collaboration with the US, Switzerland and the UK results in higher research productivity (in particular in teams from Italy, Spain, Portugal, Hungary, and Czech Republic).

> H8.2: Research teams which collaborate with research teams in the US, Switzerland and the UK have higher research productivity than teams which do not collaborate with teams from these countries.

Disciplinary versus interdisciplinary collaborations: Though we do not know of any analysis that evaluates the effects of interdisciplinary research collaboration *between* research teams on their research productivity, we expect that the positive effects prevail. However, they are certainly difficult to measure, as they might be substituted by an interdisciplinary team and collaboration *within* a research team. Therefore, in order to isolate the role of between-team collaboration the team structure has to be held constant.

> H8.3: Research teams involved in interdisciplinary research collaborations have higher research productivity than teams which do not collaborate with teams from other disciplines, provided that they have the same team structure.

Public versus private collaborations: Research collaborations with private firms create both benefits and costs: benefits are in particular the increased practical relevance of the research and additional funding, whereas the costs are due to a different orientation and logic of the partners and the danger of a lower scientific significance of the results. Hence, we expect positive effects of public-private collaboration compared to single authorships (at the level of research groups) but not compared to intra-academic collaborations.

> H8.4: Research teams involved in research collaborations with private partners have higher research productivity than teams which do not collaborate.

> H8.5: Research teams involved in research collaborations with private partners have lower research productivity than teams which collaborate with public research organisations.

Operationalisation: The collaborators of the research teams in the sample will be assessed through analysing the co-authors to their publications. By means of evaluating the address fields of these co-authors in the Web of Science, the countries of their organisations will be obtained as well as a distinction between public and private organisations. However, the disciplines of the collaboration partners cannot be detected reliably through evaluating the address information. This requires more detailed information on their participation to higher education programmes (education-based discipline) and/or research (research-based discipline). We will obtain very general measures for this in the survey.

2.4.4 Further covariates of research productivity

We discussed in section 2.3 the significance of characteristics of the parent organisation, scientific discipline and country on the research productivity of research teams. Formal characteristics like decision-making procedures and the general focus on scientific research, as well as more informal issues like the climate in an organisation and the degree of collegiality and intellectual stimulation influence research productivity. Also, research productivities vary between scientific disciplines and countries. These factors are outside of the core topics of the present analysis, but they nevertheless are taken into consideration as additional control variables.

Operationalisation:

- Characteristics of the parent organisation: The possible characteristics which might influence research productivity are nearly unlimited. For practical reasons we have to confine our efforts to those issues which might interact with the focal parameters of the analysis, namely the structure of research teams and their collaboration activities. The questionnaire therefore includes questions on the policies in regard to the composition of research groups, the procedures of and decision-making on the filling of open positions, in particular for PhD students and postdoctoral researchers. Also, we will investigate the contractual and funding regulations of post-docs. If possible, a question on particular support and integration measures for personnel from abroad should be included. Moreover, the questionnaire will try to assess, whether the environment of the research teams and their relationships to other teams of their organisation are characterised by a spirit of collegiality and cooperation.
- *Scientific discipline and field* are assessed in the survey through a question on the participation to PhD programmes (based on the ISCED 1997 classification, see annex 1) and through the bibliometric analysis (see table 1 on p. 49).
- Country of the research team

3 Research approach in work-package 1

3.1 General overview of the NetReAct approaches

NetReAct employs a methodological approach that rests on three pillars:

- a *questionnaire-based survey* targeted principally at heads of research teams to collect data on the doctoral students and post-docs at these institutions as well as further covariates which are supposed to influence research productivity,
- *bibliometric data* (publication and citation statistics) taken from the Thomson-ISI database to assess the levels of collaboration and output produced by these institutions, and
- *webometric data* (hyperlinks) collected from the internet to evaluate the position of these institutions in the life sciences networks.

The different data sets will be combined in work-package 4 to develop profiles of life sciences research teams and analyse their research productivity. Work-packages 1 and 3 mainly rest on the data to be gathered in the survey which will be described subsequently. Bibliometric and webometric methods are described in the deliverables of work-package 2.

The data collections and analyses of the NetReAct project are carried out for 10 countries as required according to the tender specifications. Several factors influenced the decision on the selection of countries: publication output in the life sciences (based on ISI data), language (Portugal instead of Greece), contrasting science policy strategies (Hungary v. Czech Republic), membership or not of the EU (Norway v. Sweden), and – of course – the tender specifications. The research population are life sciences teams in the following countries: France, Germany, the UK, Italy, Spain, Portugal, Sweden, Norway, Hungary, and Czech Republic.

The life sciences disciplines included in the analysis were originally based on two different classifications: (1) the ISCED97 classification and (2) the K.U. Leuven-IRO Subject Classification. The ISCED97 classification was intended to be used to identify the scientific disciplines of research teams and their members (heads of the teams, PhD students, post-docs) from an education-related point of view. The K.U. Leuven-IRO classification was intended to be used for identifying the research fields of the research teams in the bibliometric analysis. The K.U. Leuven-IRO classification uses a broader concept of the life sciences and accounts for the fact that life scientists do not publish only in journals of their core fields, but also in journals of related and neighbouring fields. Three major fields – Biology (Z1-Z5), Biosciences (B) and Biomedical research (R) – cover the life sciences as defined in ISCED97, as shown in the following table. The table proposes a correspondence between ISCED97 and K.U. Leuven-IRO classes. However, this is still preliminary. The exploratory interviews taught us that the ISCED 1997 classification is not really in line with the self-perception of team leaders in the life sciences and hence we favour a different approach (see page 53). If both, the ISCED 1997 and K.U. Leuven-IRO classifications have to be used, the final correspondence is to be developed on an empirical basis combining the results of the survey with the bibliometric data.

Table 1: ISCED 1997 and K.U. Leuven-IRO classifications

K.U. Leuven – IRO Subject Classification of life sciences	**ISCED 1997 42 Life Sciences**
Z BIOLOGY (ORGANISMIC & SUPRAORGANISMIC LEVEL)	biology
Z1 animal sciences	zoology, ornithology, entomology
Z2 aquatic sciences	
Z3 microbiology	microbiology, bacteriology
Z4 plant sciences	botany
Z5 pure & applied ecology	
B BIOSCIENCES	biology
B0 multidisciplinary biology	
B1 biochemistry/biophysics/molecular biology	biochemistry, biophysics
B2 cell biology	
B3 genetics & developmental biology	genetics
R BIOMEDICAL RESEARCH	
R1 anatomy & pathology	
R2 biomaterials & bioengineering	
R3 experimental/laboratory medicine	
R4 pharmacology & toxicology	toxicology
R5 physiology	

Source: NetReAct (FHSO) based on ISCED97 and K.U. Leuven-IRO.

3.2 Survey

NetReAct implements a questionnaire-based survey to collect information on doctoral students and post-docs. This survey is also used to gather further information at research team level critical for progressing understanding of key hypotheses in the analysis and modelling of WP4. This information will help resolve spurious correlations and dismiss some important rival hypotheses when it comes to assessing the impact of collaboration and other networking activities on scientific output (WP4). Available statistics on postgraduate students and post-docs are not detailed enough to meet the objectives of the project. They would pose problems of comparability, being based on varying definitions and delimitations of subject areas or aggregates of postgraduate education, e.g. not differentiating between ISCED levels 5 (master) and 6 (PhD).

3.2.1 The research population

The primary statistical unit of analysis is the university-based research team. The identification of the population of life sciences research teams drew heavily on the International Handbook of Universities (International Association of Universities

[IAU], 2003) and the internet as sources of information. The research population was assembled in three major steps of work:

1) The International Handbook of Universities was used to identify universities with teams in the life sciences.
2) From the websites of these universities we collected the names and URLs (WWW addresses) of the life sciences teams, searching for groups mentioned by the IAU and also navigating the university faculty/department structure to look for additional groups. The name and URL collection produced “hierarchical trees” which included the information on the faculties, departments, institutes and research teams.
3) The web pages of the sample teams (see below on the sampling) were submitted to a closer review when their addresses were collected. This led to the identification of further research teams which were subsequently added to the research population.

Most university websites provided information in English. However, for the majority of countries native speakers carried out additional checks of the national language websites, too. In particular in the cases of Hungary and the Czech Republic this produced additional research teams.

The search for research teams was not limited to the core life sciences faculties according to ISCED97, but life sciences teams from neighbouring faculties were also included. In particular faculties of medicine, veterinary science, and agriculture were also evaluated and their life sciences teams were added to the file.

A particular challenge was the distinction of university research teams from teams of non-university research organisations and other tertiary education institutions (colleges). In most countries the exclusion of the non-university research sector was easy to implement, because the organisations appear as separate entities. For instance the institutes of the Czech Academy of Sciences, the German Max-Planck-, Fraunhofer and Leibnitz-Societies, the Spanish and Italian Research Councils (CSIC and CNR) are in most cases not included on university websites; and if they are included, they are clearly denoted as extra-university. Things are different in France: 85% of the CNRS research teams are in cooperation with universities and other research institutions and 45% of the INSERM laboratories are located in universities; many of these labs are run under cooperation agreements or they are actually mixed laboratories, so called “Unités mixtes de recherche (UMR)”. Moreover, the temporary nature of the labs in France made the identification via the internet difficult, in particular if they were discontinued after the latest funding period. We opted to include these mixed teams, if a participation of university personnel was clearly discernible. This was applied in particular to the French case but also to other countries, when non-university and university personnel formed together a research team.

Other tertiary education institutions – in particular teaching-oriented colleges in the UK, university colleges in Sweden, and the *Fachhochschulen* in Germany – were generally not included in the research population. They are higher education institutions and in many cases they may even train PhD students, but they do not issue doctoral degrees themselves, and doctoral students play a different role than in "regular" universities. Hence, only research teams which are part of organisations issuing doctoral degrees were included in the population. The questionnaire will include a control question on this issue, in order to check whether the ex ante classification has been correct.

The research population thus identified consists of more than 6,300 teams from 356 universities (see table 2). We found more than 1,300 teams each in France and Germany and 200 or less teams in Portugal, Czech Republic, Norway, and Hungary. The average number of life sciences teams per university varies from more than 30 in the Nordic countries to only 10-11 in the UK and Portugal.

Table 2: Research population of life sciences research teams by country

Country	No. of universities with life sciences research teams	No. of life sciences teams identified	Life sciences teams per university
CR	11	160	14.5
DE	61	1,316	21.6
ES	48	835	17.4
FR	57	1,332	23.4
HU	10	214	21.4
IT	45	730	16.2
NO	6	197	32.8
PT	14	157	11.2
SE	15	505	33.7
UK	89	896	10.1
Total	**356**	**6,342**	**17.8**

Source: NetReAct (FHSO).

3.2.2 Sampling and sample characteristics

For the statistical analyses of work-package 4 an overall dataset of at least 300 research teams is desirable to secure sufficient variation. On the basis of a country sample of 10 countries this would consist of on average 30 research teams per country, perhaps 35 per large country and 25 per small country. Response rates of newer postal surveys in an academic environment measured between 25% and 50% (Barjak & Harabi, 2004; Bozeman & Corley, 2004; Fritsch & Schwirten, 2002; Laredo, 1999; Walsh, Kucker, Maloney, & Gabbay, 2000). Assuming 30% in planning the survey seems therefore adequately conservative. On this basis, the overall sample contacted would have to consist of at least 1,000 research teams.

In order to obtain a representative sample of at least 1,000 research teams we employed a stratified, random sampling which was realised in different steps. These steps are described in the following paragraphs.

1) In the first step, strata for sampling were built according to country and a simple importance indicator derived from the webometric analysis. Automatic hyperlink retrieval through the Google search engine produced the number of inlinks, or hyperlinks pointing to the homepages of the life sciences teams in the research population. An example of the search is shown in figure 1.

Figure 1: Retrieval of inlinks to the homepage of a research team with Google

Source: NetReAct (SCIT).

The link numbers obtained were used for the stratification of the population according to the presence or impact on the internet. Based on previous research (Thelwall & Harries, 2004), we assume that this presence is related to the scientific performance of the research teams. Table 3 shows the inlinks to the research teams' homepages in the 10 NetReAct countries. The average number of inlinks and the number of teams for which we couldn't find a website differ notably between the 10 countries. Teams from Southern European countries and the new member states tend to be less visible on the internet than teams from the UK, Germany and the Nordic countries. In the latter countries 10% or more of the teams – in Norway 17% – receive 10 or more inlinks to their homepage and relatively few (25-40%) do not receive any inlinks.[2]

[2] The Portuguese figures are somewhat anomalous in this regard, as more than one quarter of the units received more than 10 inlinks and the average inlink figure is 9.8 inlinks per unit. However, a closer examination reveals that the high inlink data is due to the extremely high inlink figures to the units of only one institute, namely the Instituto de Tecnologia Química e Biológica of the Universidade Nova de Lisboa. The reason for these many inlinks would have to be investigated separately, but they are entirely

Table 3: Inlinks to all life sciences research teams by country[a]

Country	Inlinks per team	No website	Percentage of teams per inlink class					
			> 10	6-10	3-5	1-2	0	Total
CR	2.1	0.7%	4.7%	7.4%	6.7%	26.8%	54.4%	100.0%
DE	5.1	4.5%	11.6%	12.6%	20.7%	27.7%	27.5%	100.0%
ES	1.5	15.7%	2.3%	4.5%	6.5%	20.4%	66.3%	100.0%
FR	1.9	8.5%	2.5%	3.2%	7.3%	27.4%	59.6%	100.0%
HU	3.0	12.0%	8.2%	4.9%	8.7%	29.3%	48.9%	100.0%
IT	1.3	15.2%	2.3%	3.9%	8.2%	19.3%	66.2%	100.0%
NO	5.8	6.6%	17.3%	6.4%	12.8%	35.3%	28.2%	100.0%
PT	9.8	10.8%	26.4%	7.1%	7.1%	15.7%	43.6%	100.0%
SE	5.0	5.7%	9.5%	4.8%	15.3%	26.7%	43.7%	100.0%
UK	8.0	5.7%	12.7%	10.2%	15.4%	26.3%	35.4%	100.0%
Total	**–**	**8.7%**	**7.7%**	**6.9%**	**12.2%**	**25.5%**	**47.7%**	**100.0%**

a The inlink counts are to some extent preliminary, as the address collection for the sample units might lead to an exclusion of erroneously included research teams (e.g. because they are not university-based, not involved in research, or not part of the life sciences).

Source: NetReAct (FHSO & SCIT).

2) As a second step we added the inlink data for each research team and ordered the teams of each country by inlink class and university. We thus obtained a list of teams with the most inlinked (= best performing) teams at the top. This served the purpose of having the best performing research teams in the life sciences included.

Table 4: Drawing probabilities by country and inlink class

Country	Drawing probability per inlink class				
	> 10	6-10	3-5	1-2	0
CR	0.900	0.810	0.729	0.656	0.590
DE	0.560	0.314	0.176	0.098	0.055
ES	0.670	0.449	0.301	0.202	0.135
FR	0.640	0.410	0.262	0.168	0.107
HU	0.850	0.723	0.614	0.522	0.444
IT	0.670	0.449	0.301	0.202	0.135
NO	0.900	0.810	0.729	0.656	0.590
PT	0.900	0.810	0.729	0.656	0.590
SE	0.670	0.449	0.301	0.202	0.135
UK	0.670	0.449	0.301	0.202	0.135

Interpretation example: Out of all research teams from Czech Republic 90% of those having more than 10 inlinks, 81% of those having 6-10 inlinks, 72.9% of those having 3-5 inlinks, 65.6% of those having 1 or 2 inlinks, and 59% of those having 0 inlinks (including those without a homepage) were included in the sample.

Source: NetReAct (FHSO).

---

internal and from other websites of the institute. Therefore, we conclude that they are mainly caused by the website structure and introduce an upward bias into the Portuguese data.

3) At the same time, some randomness was needed in order to guarantee that the data will be representative at country level. Therefore, we calculated drawing probabilities which varied according to the inlink data (see table 4). The drawing probability for the highest inlink class (> 10 inlinks) was defined in a way that secured a minimum number of 100 sample teams for the countries with small life science systems (NO, PT, HU, CZ), 125 for the countries with medium-sized systems (IT, ES, SE), and 150 teams for the countries with large systems (FR, DE, UK). Further drawing probabilities per link class are based on a simple exponential function: $dp_i = dp^i$ (dp: drawing percentage, i: inlink classes 1-5).

4) The drawing probabilities calculated in the previous step were then used to draw the sample country-wise from the sorted lists of all research teams obtained in step 2. The drawing was done systematically from the beginning of the list to its end which aimed to guarantee a maximum representation of different universities in the sample.

This resulted in the sample as shown in table 5. Overall more than 1,500 teams were selected for the sample, totalling nearly 24% of the overall research population. The coverage varies notably across countries, however. For the countries with the smaller life science systems more than 50% – in Czech Republic, Portugal and Norway nearly two third – of the research population are included. This percentage goes down to only 15.8% for France. University coverage is generally higher, of course. However, the average number of teams per university is only 5.0. Taking the expected response rate of 30%, we can expect on average only 1-2 responses per university. This will have some consequences for the required analysis at university level: it will only be possible for the few larger universities in each country, but not for all 302 universities in the sample.

Table 5: Sample of life sciences research teams per country

Country	Universities included		Teams included		Teams per university
	No.	In % of population	No.	In % of population	
CR	11	100.0%	105	65.6%	9.5
DE	56	91.8%	230	17.5%	4.1
ES	39	81.3%	143	17.1%	3.7
FR	49	86.0%	210	15.8%	4.3
HU	10	100.0%	108	50.5%	10.8
IT	38	84.4%	131	17.9%	3.4
NO	6	100.0%	129	65.5%	21.5
PT	14	100.0%	108	68.8%	7.7
SE	14	93.3%	128	25.3%	9.1
UK	65	73.0%	224	25.0%	3.4
Total	**302**	**84.8%**	**1,516**	**23.9%**	**5.0**

Source: NetReAct (FHSO).

In table 6 we provide the link statistics for the teams in the sample. As desired the average inlink figure per team is higher in the sample than in the research population (compare tables 6 and 3) which serves the purpose of having the top performers in life sciences research overrepresented. For all countries – except for Germany – the random selection of teams also led to the inclusion of teams without a homepage. Maximum percentages are 8.3% of the Portuguese teams and 7.6% of the Italian teams – this should not cause any problems for the webometric analyses of work-package 2, as the sample of available teams mainly depends on the response rate and it should still be sufficiently big.

Table 6: Inlinks to the life sciences research teams in the sample by country[a]

Country	Inlinks per team	No website	Percentage of teams per inlink class					
			> 10	6-10	3-5	1-2	0	Total
CR	2.2	1.0%	4.8%	10.6%	7.7%	26.9%	50.0%	100.0%
DE	6.0	0.0%	35.2%	21.3%	19.6%	14.3%	9.6%	100.0%
ES	4.3	4.2%	6.6%	10.2%	9.5%	19.7%	54.0%	100.0%
FR	6.0	5.2%	10.1%	8.0%	11.1%	33.7%	37.2%	100.0%
HU	4.2	1.9%	12.3%	6.6%	9.4%	28.3%	43.4%	100.0%
IT	2.8	7.6%	8.3%	9.1%	14.0%	20.7%	47.9%	100.0%
NO	6.0	3.9%	19.4%	10.5%	13.7%	32.3%	24.2%	100.0%
PT	12.6	8.3%	34.3%	8.1%	8.1%	13.1%	36.4%	100.0%
SE	11.1	1.6%	23.8%	8.7%	18.3%	21.4%	27.8%	100.0%
UK	16.2	1.8%	29.1%	16.4%	16.8%	20.9%	16.8%	100.0%
Total	8.5	3.3%	19.8%	12.0%	13.6%	22.9%	31.7%	100.0%

a The inlink counts are to some extent preliminary, as the address collection for the sample teams might lead to an exclusion of erroneously included research teams (e.g. because they are not university-based, not involved in research, or not part of the life sciences).

Source: NetReAct (FHSO & SCIT).

When the addresses for the research teams in the sample were collected from the WWW, additional information on the personnel was added. This will serve the purpose of comparing the responses to the survey with the overall sample and checking for any bias in the responses. The results are shown in figure 2 and table 7. The average size of the research team is 18.4 scientific and non-scientific workers, ranging from 13.4 in Portugal to 22 in the UK and Hungary. Across all countries PhD students represent 32.5% of the total scientific and non-scientific staff – this is in the range of the figures obtained in previous studies (Carayol & Matt, 2004; Laredo, 1999, 2001; Mangematin & Robin, 2003). PhD students are particularly relevant in Portuguese, British, Czech, and Spanish teams, whereas they are a lot less present in Hungarian and Norwegian teams. Post-docs constitute nearly one fourth of the personnel. According to this first data collection, they are more important in the UK (32%), Portugal (30%), Italy (28%), and Norway (27%) and less important in France (9%) and Hungary (13%). However, it has to be noted that the figures on PhD students and post-docs are very much preliminary and only suitable for getting a vague overview of the sample structure. They were gathered from the World Wide Web, and even for those teams which provided them, we do not know, whether they are up-to-date and complete (probably in many cases they are not).

Table 7: Average staff per life sciences research team by country

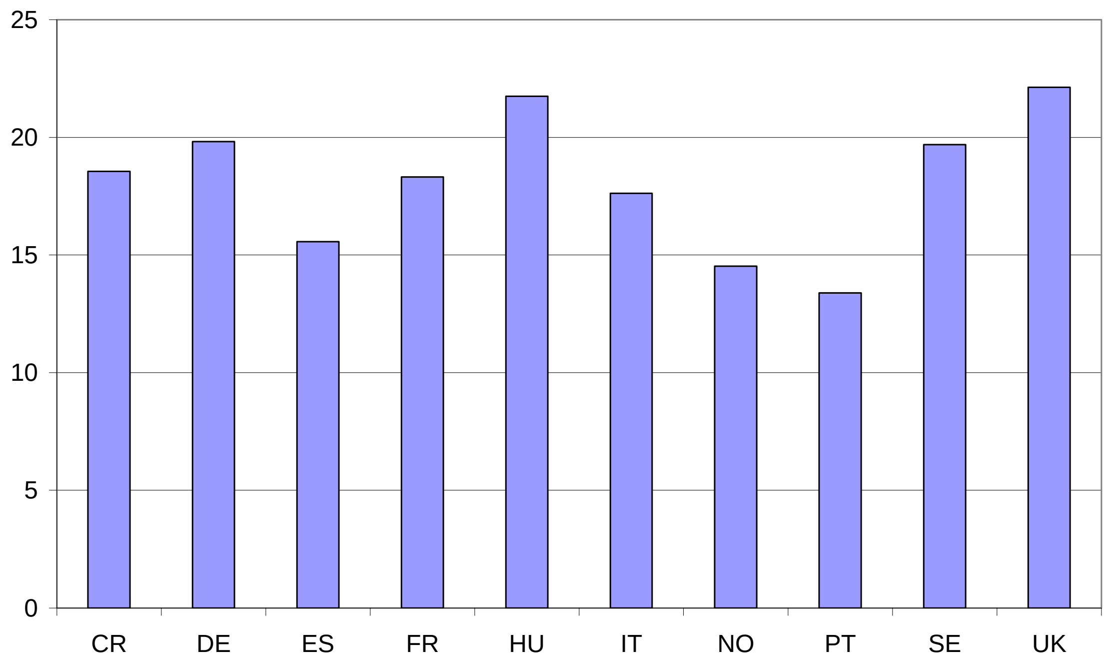
Country	Average total scientific and non-scientific staff	PhD students[a]		Post-docs[b]	
		Average no.	In % of average total staff	Average no.	In % of average total staff
CR	18.6	7.8	42.0%	4.0	21.6%
DE	19.8	6.1	30.8%	3.9	19.4%
ES	15.6	6.3	40.4%	3.4	21.7%
FR	18.3	5.0	27.4%	1.7	9.3%
HU	21.7	4.3	19.6%	2.8	12.9%
IT	17.6	4.6	26.0%	4.9	27.7%
NO	14.5	2.9	19.7%	3.9	26.9%
PT	13.4	7.1	53.3%	4.0	29.7%
SE	19.7	6.5	33.3%	4.2	21.3%
UK	22.1	9.6	43.6%	7.1	32.2%
Total	**18.4**	**6.0**	**32.5%**	**4.2**	**22.8%**

a Based on the information provided on the websites of 728 teams of the sample (= 48%).

b Based on the information provided on the websites of 567 teams of the sample (= 37%).

Source: NetReAct (FHSO).

Figure 2: Average total staff per life sciences research team by country



Source: NetReAct (FHSO).

3.2.3 Survey methodology

The survey will be carried out as an on-line survey. An introductory e-mail will be addressed to the heads of the research teams in the sample. The e-mail includes a personalised link to the website on which the questionnaire is uploaded. In order to reduce non-responses that would be caused by this methodology, additional download versions will be provided that can be filled in and returned via e-mail and/or regular mail. The questionnaire covers the key issues of NetReAct that cannot be answered through the bibliometric and webometric analyses. In particular it contains questions on:

- the research teams, their overall personnel structure, decision-making procedures, and resources,
- doctoral students and post-docs, including demographic and education-related information and recruiting strategies and procedures,
- the heads of the research teams,
- collaboration motives and trends,
- further data to help link into and evaluate the content of bibliometric and webometric databases - publication lists, web publishing strategy etc.

The questionnaire will be drawn up in English. This is not expected to have any significant negative impact on the survey, as English is the lingua franca of science and most scientists read and understand it well enough. A recent survey among scientists in five disciplines did not show any major influence of the mother tongue on the response rates to an English language questionnaire (see Barjak & Harabi, 2004).

The survey approach has been evaluated in the exploratory interviews and generally proven its feasibility. Further pre-testing of the questionnaire with an advanced draft and the same methodology and response group as in the main survey will be carried out in the near future.

3.3 Exploratory interviews

The main purpose of the exploratory interviews was to test the feasibility of certain settings of the survey, such as addressing the heads of research teams and collecting data retrospectively for the year 2003; to gather some background information on research in the life sciences and the roles of doctoral students, post-doctoral researchers and research collaborations; to test different question approaches and obtain answers for closed-ended survey questions.

The interviewees were selected according to issue-related and practical criteria: First and foremost, they should have several years of experience as active scientists and team leaders in the life sciences. Second, they should lead research teams with several PhD students and/or post-docs. Third, the interviewees should work in different countries in order to account for national particularities. Fourth, the interviewees should be within reach of the NetReAct team members and ideally existing contacts should be activated. So far, nine interviews have been carried out with life scientists who met these criteria (see annex A-2).[3] Four of the respondents are heads of research teams in France, three in the UK and two in Germany. The interviews were carried out at the end of April/beginning of May. They lasted between 40 and 120 minutes.

The interviews were planned and held as semi-structured expert interviews as described in Meuser and Nagel (1991). They were based on a detailed interview guideline with a mixture of open-ended and closed-ended questions (see annex A-3 on the guideline). This guideline included questions on the research team, the team leader, the doctoral students, post-docs and permanent scientists in the team, research collaborations, research productivity, and the role of the website. Moreover, the interviewers were asked to record different aspects of the response behaviour of the interviewees, such as misunderstandings, difficulties in giving an answer, or unwillingness to do so because of delicate questions. The interviews were either tape-recorded and transcribed, or noted down.

In the case of longer transcripts, the responses were paraphrased and condensed. They were added together with the notes to a long document that was structured along the interview guideline. In the analysis of chapter 4, the responses were compared and different meanings and interpretations were extracted. The results of this thematic comparison were verified through re-visiting the original transcripts and notes.

Chapter 4 presents a selection and interpretation of the interview results that relate to the research teams in general and to their doctoral students. The results on the other topics of the interviews (post-docs, collaboration, productivity, role of the website) will be used in the questionnaire development. Moreover, they will be re-evaluated in the making of the deliverables of the work-packages 2-4. In order to guarantee the anonymity of the respondents, the letters A to I are used instead of the interviewees' names in the text (the letters do not reflect the alphabetical order of interviewees shown in the annex table A-2).

[3] We gratefully acknowledge the willingness of all our interview partners to devote their time to the interviews, answer the questions openly and to the best of their knowledge, and support the NetReAct project. We benefited a lot from the responses and they clearly improved our understanding of research in the life sciences. We can warmly recommend all interviewees for research and policy formulation in the life sciences.

4 Interview results in regard to research teams and doctoral students in the life sciences

4.1 Research teams

4.1.1 Issues related to time and discipline of the research team

Research unit or research team?

A major problem was discovered around the term "research unit" that was originally used. The problem was mainly due to the French practice of using the word "unité" for entire institutes; however, in the German interviews questions were raised, too, on the type of organisation for which the questions were asked. Due to different organisation structures and different conventions of labelling sub-university levels, there seems to be only one possible solution: taking the lowest level of research teams which seems to be existent in all countries (according to the interviews and the information obtained from the internet in the sampling process) as the basic unit of analysis. However, to obtain meaningful data, the respondents should always answer the questions for the unit of analysis that they think is the most important in regard to the actual realisation of research. This will be the research team in most cases but not necessarily always.

Year 2003

The second problem refers to the retrospective data collection for the year 2003, though it seems not to be as grave as we feared. In order to get some correspondence with the publication data, we intend to collect most of the information within the survey for the beginning of 2003. However, this means that the respondents will have to think back two and a half years to answer the questions. Also, some of the respondents might have been affiliated to a different university or they might even have worked in a different country in 2003. The latter problem can be solved by adding questions on the research team to the questionnaire, but it will reduce the number of usable responses. The former problem of not remembering 2003 correctly certainly affects the validity of the results. One consequence for the survey is that the year should always be specified in the questions, in order to permanently remind the respondents that they have to think back in time. Particularly problematic are questions that ask for developments or time periods: in these cases, formulations like "in the past five years" will be avoided and a certain time period will be specified instead. Also, it seems impossible to collect data on earlier years than 2003.

Creation of the team

The first question of the interviews asked for the year in which the research team was founded. The oldest team was founded in 1928, the newest in December 2003. The responses point to two issues that have to be taken into account:

- Some respondents provided two dates: (1) when the team actually started to work, often when its head obtained the responsibility for two or three co-workers (PhD students and post-docs) and started more or less independent research; the teams still pertained to a larger group with a group leader who had overall responsibility; (2) when the team was officially recognised by the funding organisation. These teams were actually founded and formed by their heads.
- Some respondents joined already existing teams and didn't exactly know when these were founded. These teams were not founded by their heads; however, after obtaining the leadership, they had the possibility to influence the direction of research.

Both types of teams should be assessed in the survey, as there is clearly some relationship to the team structure and the position of the team leader.

Scientific discipline

A disciplinary classification of the research teams is required in the tender specifications and it is also desirable from an analytical perspective, for instance to evaluate the role of interdisciplinary team structures and collaborations. The tender specifications request the use of the ISCED 1997 classification that is oriented towards education and not towards research. The interviews therefore included two questions: (1) whether the team contributed to PhD programmes/education and (2) in what disciplines this would be. In some interviews, the interviewees were shown the ISCED list and asked to mark the relevant disciplines.

The responses point to several issues: Though some teams or team leaders do not have any teaching duties, they nevertheless supervise PhD students and participate in doctoral committees. This created some insecurities in regard to the correct answer on question 2 (see annex A-3) Hence, it has to be made clear in the survey that not the formal involvement in a PhD programme, but the actual supervision and training of PhD students is needed. When the ISCED 1997 list was not shown in the interview, the interviewees listed fields which are difficult to match on the ISCED classification. When the list was presented, the interviewees (1) listed several disciplines, (2) chose one or two rather general ones as the main fields, (3) sometimes stated that the question was difficult to answer, as their main field was included in various possible responses, and (4) tended to answer the question very much from a research standpoint, classifying their research results (one interviewee even included the disciplines of collaborators who contributed to his research).

What are the implications of this? The response of one interviewee might point to a possible solution: “As far as I can see, I think that genetics and biochemistry come closest to what we actually do and what people learn from us in their experimental work.” (Interviewee F, translated by FB). The main focus of the present analysis is on research. The graduate education and training activities of the life sciences teams are very much connected to their research activities. Therefore, we think it would be better to use a research-related classification instead of ISCED 1997. This should reduce the difficulties of the respondents (or later the researchers) to provide an appropriate classification of the team and it raises the validity and reliability of this information. Also, the combination of the survey and bibliometric datasets will be easier in the data analysis. For these reasons, we propose to use the K.U. Leuven – IRO Subject Classification only.

4.1.2 Personnel structure and development

Personnel of the team

The interviewees were also asked about the size and personnel structure of their teams. The responses are shown in table 8. Team sizes vary, but to some extent this also reflects the insecurity in regard to the appropriate unit of analysis. The two large teams C and D with 28 or more members consist of more than one team. However, the interviewees considered them and not the individual teams as the relevant unit of analysis.

Table 8: Size and structure of the interviewees' teams

	Total	Prof.	Other perm. faculty	Post-docs	PhD students	Masters students	Technicians	Guests and visitors	Others
A	26.5	4	4 x 20%	2	8	7	1.5	2	1
B	17.5	5	0	2	5	5	0.5 + shared staff	0	0
C	28-32	2	1	9	6-8	4-6	6	0	0
D	28	1	1	3	12	6	3	0	2
E	16	1	1	4	6	0	4	0	0
F	18	0	1	5	4	1	6	0	1
G	11	0	3	4	3	0	1	0	0
H	10-11	(0)	(3)	3	2-3	1	1	0	0
I	17.5	1	1	4	8	(1)	1.5	(1)	0

Source: NetReAct interviews.

The structure of the teams varies. In most cases there are more PhD students than post-docs, but in some cases it is also the other way around. Four issues appear from the interviews:

- Sometimes respondents provided information on part-time positions (A, B, I), but mostly they didn't. Usually, the number of persons and not the number of full-time equivalents is assumed to influence the results of research. Also, in some countries PhD students are generally given part-time positions, as they are supposed to write their theses in the rest of their working time. We will consider information on part-time positions only, when the respondents think it is relevant and provide it themselves.

- Interviewee H mentioned a particular issue related to the funding of PhD students: due to a lack of own funding two PhD students who are registered at his university carry out their doctoral work at an US American university. As these students are affiliated to a different lab and their publications also appear under a different label, we will not include any team members which are on permanent leaves.
- Master students: Seven interviewees also trained master students in their teams; but there is no general rule on whether they are counted as full team members. Interviewee G stated that, of course, he would include master students, whereas interviewee H said that he would leave them out as they are trained but do not work themselves. Hence, the survey will collect information on master students and resolve empirically whether they affect the output produced by a team.
- Visitors and guest researchers: In the interviews, we also asked for visitors and guest researchers staying for more than 12 months with a team; we learned that this is very unusual in the life sciences. More often, visitors stay for a short term of 1-3 weeks to learn a technique. In few cases longer stays of up to 2 months were mentioned. This type of guests does not stay long enough to affect the results of research. Instead, visitors may be considered as an indicator for the reputation of the research team, as obviously other scientists can learn something from it.

Policies regarding the composition of research teams

Another interview question asked for recruitment policies, e.g. gender-related, regarding the presence of scientific disciplines or certain research paradigms. Three out of the nine respondents stated that they didn't follow any policies. In these cases open positions are filled according to the applicants' excellence and fit to the position (interviewee D); one interviewee (H) said that he would appreciate a lot if he were given the power to structure the research team according to his ideas, but that this would be totally impossible, given the existing recruitment practices and procedures at his organisation. Three respondents stated that they would expect certain qualifications in regard to the field of education or knowledge of theories. Two interviewees highlighted that they would appreciate a mixture of qualifications and nationalities in their teams. Interviewee E expressed this nicely:

"I may give you an example. […] When we are working on a paper the touch and the design of an Italian is very important, […], the high technical …, the fingers of the Germans are wonderful; the critique of an English …; the kind of imagination sometimes of the French; I like also the rigueur from the Swiss people – and when you put all of that stuff together, I think it is a wonderful cocktail."

Two respondents stated that they tried to obtain a gender-balance in their teams, but one also pointed to clear barriers to this which originate in the regulations of funding (specifically the refunding of salaries paid to women on maternal leave).

The question produced some interesting responses, but it seems to be difficult to reduce the variety to valid and meaningful closed-ended responses.

Development of the team

Out of the nine teams whose leaders were interviewed, 5 teams grew, 2 remained constant and 2 decreased. Growth was above all caused by an increase of post-docs and PhD students. Some respondents stated that the development was staggered and dependent on the acquisition of grants (interviewees B and F). In the German system, the leave of professors often also entails a loss of non-permanent positions which were attached to these professors.

Table 9: Development of the interviewees' teams in the last five years

	Development	Remarks
A	Growth	all permanent staff
B	Constant	
C	Growth	
D	Decline	permanent staff and attached non-permanent positions
E	Growth	Post-docs and PhD students
F	Constant	
G	Growth	Post-docs and PhD students
H	Decline	less permanent staff, but more post-docs
I	Growth	Post-docs and PhD students

Source: NetReAct interviews.

4.1.3 Research funding

We included three different questions on the funding of the research teams which performed quite differently in the interviews:

(a) Core funding from the parent organisation (Q9): Only few respondents were able to estimate rough figures of the core funding that they received from their parent organisation. Moreover, there remain several doubts concerning the validity of this data: first, several respondents were unsure whether they should include the salaries of the permanent staff which were usually paid by the parent organisation. Second, some stated that the core funding was "lost" in the institution's general budget and that it served only to keep the infrastructure going; they didn't receive anything from this for their research, and if they did, it was impossible to state the exact amount. Third, an additional problem results from distinguishing exactly between the funding for research, teaching, and other duties of the team. Universities often do not have this information which would require detailed budgeting and controlling procedures down to the level of institutes and teams. Hence, we have to suspect that the responses to the core funding question are burdened severely by validity problems.

(b) Additional funding from other sources (Q11): Due to the problems in regard to core funding, it is almost impossible to obtain information on the teams' overall funding structure without having access to detailed budget information from the university administrations. The interviewees seemed to be somewhat more at ease with estimating the amount of additional financial support which they obtained through grants and contracts from industry. National research funding agencies, the EU framework programmes, foundations and charities, and private firms play important roles. A problem results from the varying timelines of grants and projects; also, the salaries of students and post-docs are frequently covered by studentships and fellowships. In general, it is still demanding to integrate this information and annualise the payments.

(c) Funding for PhD students and post-docs (Q10): The easiest and least complex question referred to the financing of PhD students and post-docs only. We asked whether their funding was included in the core funding, but the interviewees frequently explained from what sources their non-permanent scientific staff received their salaries. Often it was through studentships and fellowships. Only in one case (interview A) the core funding covered a major part of the expenses for PhD students and post-docs.

In our opinion, the responses given in the interviews imply that it is not possible to obtain valid information on the overall sources of financial support to research teams through a questionnaire. The funding structure is too complex to be translated into understandable questions. Hence, we will limit the financial information on the funding of post-docs and PhD students only.

4.2 Doctoral students

The interviews addressed several topics related to PhD students, in particular the recruitment process, the decision-making procedures, and the sought qualifications. Moreover, three questions dealt with the problem of exploiting PhD students and breaking the implicit contract which governs their work. Also, we asked the interview partners whether they would be able to provide the needed demographic and structural information for their PhD students in 2003. Some of the questions were asked in a similar manner for PhD students and post-docs. The present section focuses on the responses for PhD students; the responses on post-docs will be taken into account in the future work of work-package 3.

4.2.1 Recruitment of doctoral students

4.2.1.1 Recruitment processes

The question on the recruitment process for PhD students points to some country differences which originate in differing education systems. We can summarise these differences and distinguish three models:

There is first the “graduate school model”. A graduate school is established with a particular research programme, often as a collaborative programme of several professors and institutes. In this case, doctoral students are usually sought after the graduate school has been established and funding has been secured. Students are recruited nationally or even internationally through ads in the different available media, in particular the internet, newspapers and scholarly journals. Moreover, the participating professors spread the word among their colleagues through e-mail and messages to mailing lists and blackboards. The applicants are evaluated by a jury that consists of representatives from the participating organisations. The graduate school model is predominant in the Anglo-American higher education system but it has spread to other countries in the meantime (e.g. the German interviewees also pointed to graduate schools).

The other two recruitment models are more individualised than the graduate model and the acceptance of PhD candidates depends more on the individual professor and team leader. They also differ in regard to the starting point and the sources of funding:

New grants might ask for new qualifications in a team or include studentships for which new team members are sought. We call this recruitment model "job competition model". Grant-based and the newly created PhD positions have to be filled through advertising and informing colleagues in a similar way as in the graduate school model. Additionally, the team leaders might recruit their own graduates or select candidates from pools of unsolicited applications. The recruitment usually is carried out by the principal investigator who obtained the grant. It involves the team, if the applicants are presented and seminars and discussions with them are organised. Other university posts or bodies are involved only in regard to formal issues, i.e. through checking whether the applicants meet the formal requirements of the examination regulations.

The third model roots in different types of job applications and we can call it the "application model". It is vice versa to the job competition model, as the team leader first has a PhD student and then tries to obtain appropriate funding. There are several variants to this model:

- The most favoured solution is when PhD students already have a studentship and offer their work at lowest additional costs to their team leader and supervisor to-be.
- In case of good and promising applicants or masters graduates, the team leader might decide to support the student's application for a studentship; or he might even look for other funds himself, e.g. grants and contracts, which could be used for funding the student.
- The French system used to be somewhat specific, as the doctoral students to-be entered a competition for studentships. The lab obtained a PhD student within this regular funding scheme only, if they succeeded in this competition. Team leaders complained of their lack of influence on this selection process. However, it is obviously being reformed currently.

Table 10: Recruitment models for doctoral students

	Graduate school model	Job competition model	Application model
A	Yes	Yes	Yes
B		Yes	
C	Yes	Yes	Yes
D	Yes		Yes
E		Yes	Yes
F		Yes	Yes
G		Yes	Yes
H		Yes	Yes
I	Yes	Yes	Yes

Source: NetReAct interviews.

In particular the job competition model and the application model seem to be present nearly everywhere. However, their relevance might differ between the research teams and the questionnaire should therefore try to assess the frequency of the different models in 2003.

Leadership and decision-making

As stated in section 2.1.3 (page 26), the practices of decision-making also influence the productivity of life sciences teams. Hence, we included a brief question on the general management practices and additional questions for the making of decisions on the recruitment of PhD students and post-docs.

The responses to the general question point to an accentuated position of the team leaders in all cases. However, all interviewees also mentioned the involvement of other team members, e.g. the most experienced scientists, the leaders of sub-groups or even the entire team, in discussions that lead to decisions. Hence, it seems to be difficult to assess the actual differences in regard to the decision-making process with this kind of rather superficial question directed at the team leaders only. Other team members would have to be interviewed, too.

The process of accepting applicants for PhD positions doesn’t seem to differ from this general decision-making, except for the case of graduate schools. When PhD positions are filled outside of graduate schools, the team leader is again the main decision-maker, but he usually involves his experienced team members and those with whom the applicant would have to work afterwards to get a second opinion on promising applicants.

Important factors in the recruitment process

Several different factors which influence the recruitment process were listed in the interviews:

- Qualifications: Nearly every interviewee listed the qualifications of the applicants, their areas of expertise, experimental skills and knowledge of techniques as important criteria in the acceptance of an applicant. Interviewee E stressed that he would not look only at the applicants' qualifications from an isolated perspective, but also at a complementary relationship to existing qualifications in the team.
- Research skills: In addition to the formal and technical skills, the ability to work independently and without constant supervision was also mentioned as an important asset; another sought skill relates to critique, independent thinking, and reflection.
- Applicants from certain countries and internationally mobile applicants also received a bonus (interviewees F and G).
- Motivation and interest in the work of the team were also stated as important criteria.
- Last but not least, the social competence of the applicants and whether they seemed to match with the other members of the team were mentioned.

Most interviewees also stated that they usually undertake the effort of evaluating the information that they obtain from the applicants; these evaluations can use different criteria and methods: (1) the applicants' grades, (2) publications and their impact factors (however, this is more relevant for post-docs, as graduates usually have not yet published), (3) recommendations from previous supervisors, (4) their knowledge of the field and those who work in it – as interviewee C stated: "one knows the people that come from there".

4.2.2 Fulfilment of the implicit contract

The implicit contract between the research team leaders and their PhD students has been described as an important element that governs the role, impact, and future career opportunities of PhD students (Mangematin, 2000; Stephan & Levin, 1997, 2001a). When it is broken systematically, the costs and benefits of a PhD degree become unbalanced and a research team becomes less attractive for PhD students. We collected information on the fulfilment of this implicit contract and on the role of doctoral students in research teams through three specific questions.

Functions of PhD students

We can assume that the main objective of PhD students is to write their thesis. In doing so, they create new scientific knowledge and raise their job prospects. Any other obligation, like doing experimental work for their principal investigator, taking on teaching duties, or doing technical and administrative work can keep them from reaching their main objective. Three out of seven interviewees who responded to this question stated that the PhD-related research was the most important objective of PhD students. In several cases the interviewees listed a mixture of functions and responsibilities: e.g. the doctoral thesis is the main project, but PhD students also should help in the training of lab students, give seminars and participate in the scientific discussion of the lab (interviewee G); PhD students work primarily on their own research; but they also organise group meetings and seminars and the caring of the lab animals, maintain the research team's website, help with the supervision of undergraduate research projects, and collaborate on grant proposals (interviewee I). In order to interpret the results, more information on the actual daily practice would be needed and what tasks take the majority of the students' time. Also, we would need to know more about the students and their own plans and activities, e.g. whether they plan to stay in science or prefer to move on to engineering and technological development outside of public research. Hence, the question provides some information on the PhD students work and tasks, but this information will possibly not be sufficient to find out, whether and to what extent the implicit contract is broken. Moreover, two interviewees seemed to be perplexed by this question, possibly because they suspected something doubtful behind an answer that was self-evident in their opinion.

Practices of exploiting PhD students

In a further question the interviewees were asked whether they knew of any practices which could be interpreted as an exploitation of PhD students, such as considering them as cheap labour. Though they were asked this question at a general level and in regard to their several years of work experience as life scientists and team leaders, some respondents took the question personal and denied the existence of such problems for their team. The question seems to be rather delicate. Three respondents said that they wouldn't know of any practices which could be classified as a misuse of PhD students. Moreover, this would not be in the interest of the team leader who also benefited from successful PhDs (interviewee G).

However, also different problems were described which at least hint at a lack of care and supervision if not exploitation: First, large-scale projects might produce scientific value only for the principal investigator; the value for the individual scientists is rather small, if the tasks are only demanding at the technical level, but not at the intellectual and scientific level (interviewee F). Second, due to a lack of competence of the team, students might not be integrated and fail to learn the scientific context of their work (interviewee G). Some PIs don't reserve enough time for the supervision of their students, in particular if these require more efforts than usual (interviewees A and E). Third, students sometimes are paid very little and this creates a feeling of being exploited by somebody who also benefits from their work but earns much more (interviewee H).

Though the responses provided some interesting insights into the problems and practices of life sciences research, the question is of limited use. Above all, it seems difficult to ask the heads of research teams such a question. In order to obtain a true picture, we would need to get the opinions of the PhD students, too.

Further career of the PhDs

A third possibility of assessing the fulfilment of the implicit contract between supervisors and PhD students could be based on the positions which they obtained after finishing their PhD. The interviewees in most cases readily knew where their former students went after graduating (see table 11). In most cases they knew the type of position/sector and the country where the graduates went.

Based on this information a typology can be constructed which could help to evaluate, to what extent the PhD education contributed to augmenting the human capital of the teams' PhD students. This in turn affects the attractiveness of a team for students. A PhD above all is needed – and therefore provides maximum value – for a career in academia and other public research; it is of less value in private business research and if the PhD works in an entirely different area; and it is useless or even counterproductive, if the PhD is unemployed (Mangematin, 2000; Stephan & Levin, 1997). Also, we might assume that positions in countries with a developed academic environment in the life sciences or a developed biotechnological and medical industry are more attractive to the graduates, as they increase their job opportunities afterwards.

Table 11: Positions of PhDs after finishing their theses

Team	PhD1	PhD2	PhD3	PhD4	PhD5
A	University, UK	University, UK	University, UK		
B	Lecturer, China	Research director, China	Scientist, China	Post-doc, own group	
C	Post-doc, own university	Post-doc, USA	Post-doc, USA	Post-doc, NL	Post-doc, DE
D	–	–	–	–	–
E	Post-doc, USA	Private firm, CH	Post-doc, DE	Post-doc	Patent Office, DE
F	Scientist, USA	Scientist, FR	Private firm, FR		
G	Post-doc, own group	Post-doc, own group	Private firm, CH		
H	Post-doc, USA	Post-doc, DK			
I	Post-doc, own university	Post-doc, Brazil	Post-doc, Argentina	Journalist, NZ	Private firm, UK

Source: NetReAct interviews.

The ranking scheme still has to be fully developed. The easiness with which the necessary information could be collected and the great variety encountered in the responses in 8 interviews convinces us that it is worth the effort. Two weaknesses also have to be mentioned, however: We do not know to what extent the responses contain a bias to the successful graduates and the unsuccessful are simply "forgotten". This problem should be even greater, if a positive selection is more or less supported through limiting the number of PhD graduates for which the information is sought. Half of the respondents in the interviews stated that they had more than 5 PhD graduates in the last five years. Hence, it seems worthwhile to collect this information on more, possibly even twice as many PhD graduates. We also did not yet include PhD students who abandoned their education before finishing it. Several reasons on both the student's and the supervisor's side might be responsible for such a decision. But all in all, it points to a negative ending of a PhD project and wasted time for the team and the unsuccessful PhD candidate.

4.2.3 Demographic and education-related information

In order to collect the information on the PhD students asked for in the tender specifications – age, country of origin, gender, field of study, university of master and duration and origin of funding – we showed the interviewees tables and asked them whether they would be able to fill this in for the students in their team in the year 2003, if they received it by e-mail (see annex A-3).

Six of the eight interviewees who answered this question stated that they could provide the needed information. One respondent said that his secretary could do it. Interviewee F stated that he would not be able to do it without consulting his records and documents on the students who were affiliated to his team. However, the interviewee also stated that (part of) this information only interests him when an open position is filled and that he doesn't think that it is relevant for the work. This might indicate some unwillingness to answer irrelevant questions which more or less is a general weakness of the survey approach and probably cannot be avoided.

All in all, the exploratory interviews so far fully met the purpose of providing additional information on the life sciences and evaluating the general approach as well as individual questions with experienced team leaders from the life sciences.

5 Outlook for further work to be done in WP 1

The further work in work-package 1 centres on two issues: (1) the preparation and realisation of the data collection on doctoral students and the analysis of the collected data; (2) the coordination with the work done in WPs 2 and 3 and preparation of the more detailed analysis of work-package 4.

Data collection and analysis for doctoral students

Several steps are still necessary in regard to the preparation and realisation of the survey of WP 1 (probably to be combined with the WP 3 survey on post-docs):

(1) Refinement of the sample and finalisation of the address collection: Currently the sample consists of 1,516 research teams. However, for some of these teams, the first attempt at collecting the necessary information on their team leaders and e-mail addresses failed. Different reasons caused this failure, for instance the supposed team turned out to be a larger department consisting of several teams, the research assistants collecting the data were not sure whether the team really belonged to a university and was active in life sciences research, or the names and addresses of the team leaders were simply not found on the internet. A senior researcher will check the information collected so far and fill in the missing information for these teams, replace them or exclude them from the sample.

(2) Development of the questionnaire: In parallel, the guideline for the exploratory interviews will be converted into the first draft of the questionnaire to be used in the survey. The results of the interviews will lead to considerable modifications, deletions and enhancements. Also, for the open-ended questions answers options and scales will be developed. As the survey is currently planned to be conducted online, the questionnaire will also have to be converted into an html-document.

(3) Pretesting the questionnaire: In order to secure an acceptable response rate in the survey, the questionnaire will be pretested with selected respondents from the target population. If possible, additional expert interviews with team leaders in the life sciences will be carried out to check early versions of the questionnaire. However, the pretest rests mainly on a fully-fledged survey of a small part of the study population. For this purpose, a random selection of 10-15% of all research teams in the sample from the large countries Germany, France, and the UK will be made (adding up to 5% of the overall sample). These teams will receive a next to final version of the questionnaire and additional questions on their opinion to the questionnaire and possible improvements.

(4) Realisation of the survey: After the pretesting of the questionnaire and the revision of the sample will have been completed, the survey will take place. The start of the survey is planned for July 1st, in order to minimise problems resulting from the upcoming holiday season. Respondents will be asked to fill in the survey within a period of four weeks, and after this time period a reminder e-mail will be sent out. The data entry is realised automatically and the analysis period will start as soon as a

sufficient number of responses will become available. We expect that first analyses will become possible in project month 7 (mid-August) which gives us two months of analysis before the deliverable 1.2 has to be delivered in mid-October.

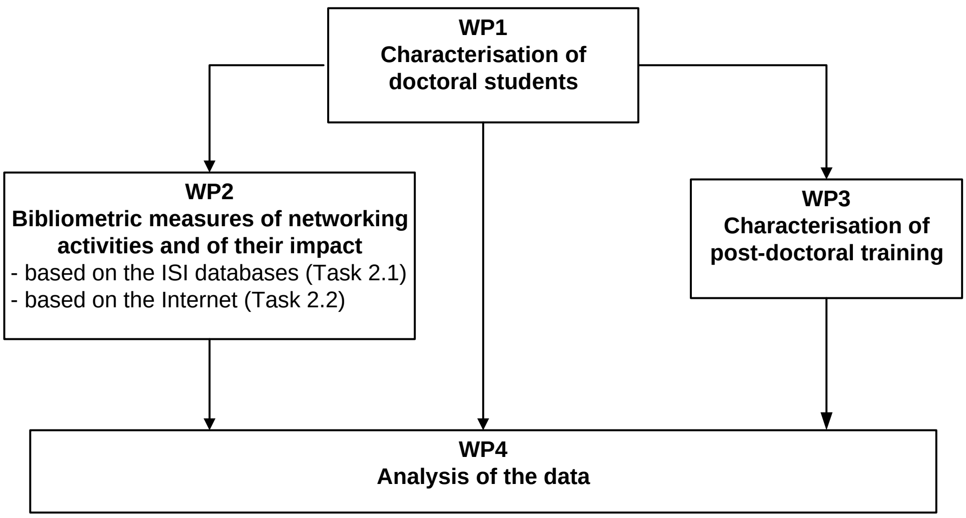
(5) Data analysis: The collected data will be analysed mainly with descriptive methods. The analysis in WP 1 will focus on the structure of research teams and the structural characteristics of the doctoral students. Questions that will be addressed are for instance: How many doctoral students work on average at organisations in the life sciences? What are the funding patterns of doctoral studies? Where do doctoral students come from and where do they go to? What relationship exists between these patterns and demographic characteristics such as age, gender, or country of birth? What strategies are apparent in the recruitment of doctoral students (e.g. in regard to their previous university, specific knowledge, firm contacts, etc.)? The analysis will be mainly carried out at team level. However, whenever feasible the data will also be aggregated to university and country levels.

(6) Methodological generalisation: The NetReAct project has the character of a pilot project in regard to the collection of information on research teams, doctoral students, post-docs and research collaborations. Therefore, one of the tasks of WP 1 will also be to synthesise the lessons learned in the data collection and discuss a possible extension to other disciplines and countries.

Coordination with the other work-packages

The second major task of work-package 1 is to secure that its results interlink with the results produced in the other work-packages of the NetReAct project. As we already pointed out in the project proposal, WP 1 provides inputs to all the other work-packages (see figure 3).

Figure 3: Interdependencies between the work-packages



Source: NetReAct.

The work in WP 2 has to produce data on the same research teams that responded to the survey to permit an overall analysis in WP 4. The coordination between WP 1 and 3 is unproblematic, as the methodology is identical. The coordination with WP 2 is more demanding.

In order to facilitate the retrieval of bibliometric data on the research teams from the Web of Science, name lists of the team members would be helpful, as the address information included in the Science Citation Index Expanded (SCIE) is not sufficiently detailed to identify all team members and their publications. Moreover, in addition to the publication and citation data, some of the data on research collaborations will also be taken from the SCIE: e.g. the total numbers, countries, and affiliations of the collaborators. However, other information such as the discipline of the collaboration partners cannot be obtained. Hence, a coordination of the data collection in WP 1/3 and WP 2 is necessary to secure that all the needed information will be available in the end.

Moreover, the webometric data collection provides data on the web presentations of the research teams and the hyperlinks which point to them. It will also need the list of teams which answered in the survey.

References

Allison, P. D., & Long, J. S. (1987). Interuniversity mobility of academic scientists. *American Sociological Review, 52*, 643-652.

Allison, P. D., & Long, J. S. (1990). Departmental effects on scientific productivity. *American Sociological Review, 55*, 469-478.

Andrews, F. M. (1979). The international study: its data sources and measurement procedures. In F.M. Andrews (Ed.), *Scientific productivity. The effectiveness of research groups in six countries* (pp. 17-52). Cambridge: Cambridge University Press.

Aran, L., & Ben-David, J. (1968). Socialization and career patterns as determinants of productivity of medical researchers. *Journal of health and social behaviour, 9,* 3-15.

Arora, A., David, P. A., & Gambardella, A. (1998). Reputation and competence in publicly funded science: estimating the effects on research group productivity. *Annales d'Economie et de Statistique 49/50,* 163-198.

Arundel, A., & Geuna, A. (2001). *Does Proximity Matter for Knowledge Transfer from Public Institutes and Universities to Firms?* SPRU Electronic Working Papers Series, no. 73. Retrieved July 16, 2003 from: http://www.sussex.ac.uk/spru/publications/imprint/sewps/sewp73/sewp73.pdf.

Baird, L. L. (1986). What characterizes a productive research department? *Research in Higher Education, 25*, 211-225.

Barjak, F. (2005). Research productivity in the internet era. Proceedings of the ISSI 2005 conference, Stockholm.

Barjak, F., & Harabi, N. (2004). *Internet for research. SIBIS topic report no. 2 to the European Commission, DG Information Society*. Retrieved January 28, 2005, from: http://www.fhso.ch/pdf/unternehmer/sibis_final.pdf

Beaver, D. deB. (2001). Reflections on Scientific Collaboration (and its study): Past, Present, and Future. *Scientometrics, 52,* 365-377.

Björneborn, L., & Ingwersen, P. (2004). Towards a basic framework for webometrics. *Journal of the American Society for Information Science and Technology, 55*, 1216-1227.

Blackburn, R.T., Behymer, C.E., & Hall, D. E. (1978). Research Note: Correlates of Faculty Publications. *Sociology of Education, 51,* 132-141.

Bland, C. J., & Ruffin, M. T. (1992). Characteristics of a productive research environment: literature review. *Academic Medicine, 67,* 385-397.

Bonaccorsi, A., & Daraio, C. (2003). Age effects in scientific productivity - The case of the Italian National Research Council (CNR). *Scientometrics, 58,* 49-90.

Bordons, M., & Zulueta, M. A. (1997). Comparison of research team activity in two biomedical fields. *Scientometrics, 40,* 423-436.

Borgman, C., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology 36*, 3-72.

Bozeman, B. (2000). Technology transfer and public policy: a review of research and theory. *Research Policy, 29,* 627-655.

Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy, 33*, 599-616.

Braun, T., Glänzel, W., & Schubert, A. (2001). Publication and cooperation patterns of the authors of neuroscience journals. *Scientometrics, 51,* 499-510.

Brooks, H. (1994). The relationship between science and technology. *Research Policy, 23,* 477-486.

Carayol, N., & Matt, M. (2004). Does research organization influence academic production? Laboratory level evidence from a large European university. *Research Policy, 33,* 1081-1102.

Charles, D., & Howells, J. (1992): *Technology transfer in Europe. Public and private networks.* London & New York: Belhaven Press.

Cohen, J. E. (1981). Publication rate as a function of laboratory size in 3 biomedical-research institutions. *Scientometrics, 3,* 467-487.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and Impacts: Survey Results on the Influence of Public Research on Industrial R&D. *Management Science, 48,* 1-23.

Cole, J.R., & Cole, S. (1973). *Social stratification in science.* University of Chicago Press: Chicago & London.

Crane, D. (1972). *Invisible colleges. Diffusion of knowledge in scientific communities.* University of Chicago Press: Chicago & London.

Cronin, B. (1982). Invisible colleges and information transfer. A review and commentary with particular reference to the social sciences. *Journal of Documentation, 38,* 212-236.

Cronin, B. (1984). *The citation process.* London: Taylor Graham.

Dasgupta, P., & David, P. (1994). Toward a new economics of science. *Research Policy*, 23, 487–521.

Davenport, S. (2004). Panic and panacea: brain drain and science and technology human capital policy. *Research Policy, 33,* 617-630.

David, P. A., & Foray, D. (2002): *Economic Fundamentals of the Knowledge Society.* Stanford Institute for Economic Policy Research, Discussion Paper No. 01-14.

David, P. A., & Keely, L. (2002). *The economics of scientific research coalitions: collaborative network formation in the presence of multiple funding agencies*. Retrieved June 08, 2003 from: http://netec.mcc.ac.uk/WoPEc/data/Papers/attwimass200211.html.

Debackere, K., & Rappa, M.A. (1994). Technological communities and the diffusion of knowledge: a replication and validation. *R&D Management, 24,* 355-371.

Etzkowitz, H., & Leydesdorff, L. (Eds.)(1997). *Universities and the Global Knowledge Economy: A Triple Helix of University–Industry–Government Relations.* London: Cassell Academic.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university-industry-government relations. *Research Policy, 29,* 109-123.

European Commission (2003). *Third European Report on Science & Technology Indicators 2003 - Towards a knowledge-based economy.* Brussels: European Commission.

Felderer, B., & Obersteiner, M. (1999). *Efficiency and Economies of Scale in Academic Knowledge Production.* Institute for Advanced Studies, Vienna, Economics Series No. 63. Retrieved March 07, 2005, from: http://www.ihs.ac.at/publications/eco/es-63.pdf

Fox, M. F. (1983). Publication productivity among scientists: a critical review. *Social Studies of Science, 13,* 283-305.

Fritsch, M., & Schwirten, C. (1999). Enterprise-university co-operation and the role of public research institutions in regional innovation systems. *Industry and Innovation, 6,* 69-83.

Fritsch, M., & Schwirten, C. (2002). R&D cooperation between public research institutions - magnitude, motives and spatial dimension. In L. Schätzl & J. Revilla Diez (Eds.), *Technological Change and Regional Development in Europe.* Heidelberg & New York: Springer.

Fry, J. (2004): The Cultural Shaping of ICTs within Academic Fields: Corpus-based Linguistics as a Case Study. *Literary and Linguistic Computing, 19,* 303-319.

Funtowicz, S.O., & Ravetz, J. (1993). Science in the post-normal age. *Futures, 25,* 739-756.

Georghiou, L. (1998). Global cooperation in research. *Research Policy, 27,* 611-626.

Gibbons, M., & Johnston, R. (1974). The roles of science in technological innovation. *Research Policy, 3,* 220-242.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies.* London, Thousand Oaks, New Delhi: Sage.

Glänzel, W., & Schubert A. (2001), Double Effort = Double Impact? A Critical View at International Co-authorship in Chemistry, Scientometrics, 50 (2), 199-214.

Glänzel, W. (2001). National Characteristics in International Scientific Co-authorship Relations. *Scientometrics, 51,* 69-115.

Glänzel, W. (2002). Co-authorship patterns and trends in the sciences (1980-1998). A bibliometric study with implications for database indexing and search strategies. *Library Trends, 50,* 461-473.

Glasner, P. (1996). From community to 'collaboratory'? The Human Genome Project and the changing culture of science. *Science and Public Policy, 23*, 109-116.

Gupta, B. M., & Karisiddappa, C. R. (1998). Collaboration in Theoretical Population-Genetics Speciality. *Scientometrics, 42,* 349-376.

Hagstrom, W. O. (1965). *The scientific community.* New York: Basic books.

Heimeriks, G., Hörlesberger, M., & van den Besselaar, P. (2003). Mapping communication and collaboration in heterogeneous research networks. *Scientometrics, 58*, 391-413.

Hemlin, S., & Gustafsson, M. (1996). Research production in the arts and humanities - A questionnaire study of factors influencing research performance. *Scientometrics, 37,* 417-432.

Hemptinne, Y. de, Andrews, F. M. (1979). The international comparative study on the organization and performance of research teams: an overview. In F.M. Andrews (Ed.), *Scientific productivity. The effectiveness of research groups in six countries* (pp. 3-15). Cambridge: Cambridge University Press.

Hicks, D. (2004). The four literatures of social sciences. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 473-495). Dordrecht, Boston, & London: Kluwer.

Hicks, D., & Hamilton, K. (1999). Does university-industry collaboration adversely affect university research? *Issues in Science and Technology, 15*, 74-75.

Hilgartner, S. (1995). Biomolecular database - new communication regimes for biology? *Science Communication, 17,* 240-263.

Hollingsworth, J. R. (2002). *Research organizations and major discoveries in twenthieth-century science: a case study of excellence in biomedical research.* WZB discussion paper P02-003. Retrieved March 08, 2005 from Wissenschaftszentrum Berlin für Sozialforschung: http://skylla.wz-berlin.de/pdf/2002/p02-003.pdf.

Hollingsworth, J. R., & Hollingsworth, E.J. (2000). *Major Discoveries and Biomedical Research Organizations: Perspectives on Interdisciplinarity, Nurturing Leadership, and Integrated Structure and Cultures.* Retrieved March 08, 2005 from http://www.umu.se/inforsk/universitets ligan/hollingsworth.html.

International Association of Universities (2003). *International Handbook of Universities.* 17th edition, Paris: IAU.

Johnston, R. (1994). Effects of resource concentration on research performance. *Higher Education, 28*, 25-37.

Joly, P.B., & Mangematin, V. (1996). Profile of laboratories, industrial partnerships and organization of R&D: the dynamics of relations with industry in a large research organization. *Research Policy, 25,* 901-922.

Katz, S.J. (1994). Geographical proximity and scientific collaboration. *Scientometrics, 31*, 31-43.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy, 26*, 1-18.

Kline, S. J., & Rosenberg, N. (1986). An overview of innovation. In R. Landau & N. Rosenberg (Eds.): *The positive sum strategy* (pp. 275-305). Washington D.C.: National Academy Press.

Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences and the Shaping of Electronic Media in Supporting Scientific Communication. *Journal of the American Society for Information Science, 51,* 1306-1320.

Knorr, K. D., Mittermeir, R., Aichholzer, G., & Waller, G. (1979a). Individual publication productivity as a social position effect in academic and industrial research teams. In F.M. Andrews (Ed.), *Scientific productivity. The effectiveness of research groups in six countries* (pp. 55-94). Cambridge: Cambridge University Press.

Knorr, K. D., Mittermeir, R., Aichholzer, G., & Waller, G. (1979b). Leadership and group performance: a positive relationship in academic research teams. In F.M. Andrews (Ed.), *Scientific productivity. The effectiveness of research groups in six countries* (pp. 95-120). Cambridge: Cambridge University Press.

Knorr-Cetina, K. (1981). *The Manufacture of Knowledge. An Essay on the Constructivist and Contextual Nature of Science.* Oxford: Pergamon Press.

Knorr-Cetina, K. (1983). The ethnographic study of scientific work: towards a constructivistic interpretation of science. In K. Knorr-Cetina & M. Mulkay (Eds.), *Science Observed. Perspectives on the Social Studies of Science* (pp. 115-140). London: Sage.

Kraut, R.E., Egido, C., & Galegher, J. (1990). Patterns of contact and communication in scientific research collaboration. In R.E. Kraut, C. Egido, & J. Galegher (Eds.), *Intellectual teamwork. Social and technological foundations of cooperative work* (pp. 149-171). Hillsdale: Lawrence Erlbaum Associates.

Larédo, P. (1999). The Development of a Reproducible Method for the Characterisation of a Large Set of Research Collectives. A Test on Human Genetics in Europe. A report of the PSR project of the EU TSER programme. Paris: CSI. Retrieved April 4, 2005, from: http://www.sussex.ac.uk/Units/spru/biotechnology/hgwhole.pdf.

Larédo, P. (2001). Benchmarking of RTD policies in Europe: "Research Collectives" as an entry point for renewed comparative analyses. *Science and Public Policy, 28*, 285-294.

Larédo, P., & Mustar, P. (2000). Laboratory Activity Profiles: An Exploratory Approach. *Scientometrics, 47,* 515-539.

Laudel, G. (2001). Collaboration, creativity and rewards: why and how scientists collaborate. *International Journal of Technology Management, 22,* 762-781.

Laudel, G. (2003). Studying the brain drain: Can bibliometric methods help? *Scientometrics, 57,* 215-237.

Lee, Y. S. (1996). 'Technology transfer' and the research university: A search for the boundaries of University Industry Collaboration. *Research Policy, 25,* 843-863.

Leeuwen, T. N. van, Moed, H. F., Tijssen, R. J. W., Visser, M. S., & Raan, A. F. J. van (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics, 51*, 335-346.

Levin S. G., & Stephan P. E. (1991). Research Productivity over the Life Cycle: Evidence for Academic Scientists. *American Economic Review, 81,* 114-132.

Liebeskind, J. P., Oliver, A. L., Zucker, L., & Brewer, M. (1996). Social networks, learning, and flexibility: Sourcing scientific knowledge in new biotechnology firms. *Organization Science, 7,* 428-443.

Lievrouw, L. (1988). Four programs of research in scientific communication. *Knowledge in Society, 1*(2), 6-22.

Long, J.S. (1978). Productivity and academic position in the scientific career. *American Sociological Review, 43,* 889-908.

Long, J. S., & McGinnis, R. (1981). Organizational context and scientific productivity. *American Sociological Review, 46,* 422-442.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences, 16,* 317-323.

Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology, & Human Values, 17,* 101-126.

Mangematin, V. (2000). PhD job market: professional trajectories and incentives during the PhD. *Research Policy, 29,* 741-756.

Mangematin, V., & Robin, S. (2003): The Double face of PhD students. The Example of Life Sciences in France. In J. Enders, & E. Weert, (Eds.), *Proceedings of the International workshop Science, Training and Career. Changing modes of knowledge production and labor markets* (pp. 109-124). Center for Higher Education Policy Studies: Enschede.

Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy, 20,* 1-12.

Mansfield, E. (1998). Academic research and industrial innovation: an update of empirical findings. *Research Policy, 26,* 773-776.

Melin, G., & Persson, O. (1996). Studying research collaboration using co-authorships, *Scientometrics, 36*, 363-377.

Meyer-Krahmer, F., & Schmoch, U. (1998). Science-based technologies: University-industry interactions in four fields. *Research Policy, 27,* 835-851.

Mulkay, M. (1977). Sociology of the scientific research community. In Ina Spiegel-Rösing & Derek de Solla Price (Eds.), *Science, technology and society. A cross-disciplinary perspective* (pp. 93-148). London & Beverly Hills: Sage.

Mulkay, M.J., Gilbert, G.N., & Woolgar, S. (1975). Problem areas and research networks in science. *Sociology, 9,* 187-203.

Murray, F. (2002). Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Research Policy, 31,* 1389-1403.

Narin, F., Stevens, K., & Whitlow, E. S. (1991). Scientific collaboration in Europe and the citation of multinationally authored papers. *Scientometrics, 21,* 313-323.

National Science Board (2002): *Science & engineering indicators 2002.* Arlington, Va.: National Science Foundation. Retrieved April 03, 2004, from: http://www.nsf.gov/sbe/srs/seind02/pdfstart.htm.

National Science Board (2004): *Science & engineering indicators 2004.* Arlington, Va.: National Science Foundation. Retrieved September 29, 2004, from: http://www.nsf.gov/sbe/srs/seind04/start.htm.

Newman, M.E.J. (2000): *The structure of scientific collaboration networks.* Santa Fe Institute WP 00-07-037. Retrieved August 13, 2003 from: http://www.santafe.edu/sfi/publications/Working-Papers/00-07-037.pdf.

Nissani, M. (1997). Ten cheers for interdisciplinarity: the case for interdisciplinary knowledge and research. *Social Science Journal, 34*(2), 201-216.

OECD (2001). *Towards a knowledge-based economy. STI Scoreboard 2001*. Paris: OECD.

Pavitt, K. (1998). The social shaping of the national science base. *Research Policy, 27,* 793-805.

Persson O., Glänzel W., & Danell R. (2004), Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies, Scientometrics, 60 (3), 421-432.

Pravdic, N., & Oluic-Vukovic, V. (1986). Dual approach to multiple authorship in the study of collaborator/scientific output relationship. *Scientometrics, 20,* 235-255.

Price, D. J. de Solla, & Beaver, D. deB. (1965). Collaboration in an invisible college. *American Psychologist, 21,* 1011-1018.

Quin, J., Lancaster, F.W., & Allen, B. (1997). Types and levels of collaboration in interdisciplinary research in the sciences. *Journal of the American Society for Information Science, 48,* 893-916.

Raan, A. F. J. van (2003). The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschätzung - Theorie und Praxis, 12*(1), 20-29.

Raan, A. F. J. van (2004). Measuring science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 19-50). Dordrecht, Boston, & London: Kluwer.

Rahm, D. (1994). Academic Perceptions of University-Firm Technology Transfer. *Policy Studies Journal, 22,* 267-278.

Ramesh Babu, A., & Singh, Y.P. (1998). Determinants of research productivity. *Scientometrics, 43,* 309-329.

Ramsden, P. (1994). Describing and explaining research productivity. *Higher Education, 28,* 207-226.

Rappa, M.A., & Debackere, K. (1992). Technological communities and the diffusion of knowledge. *R&D Management, 22,* 209-220.

Rogers, E.M. (2002). The Nature of Technology Transfer. *Science Communication, 23,* 323-341.

Salter, A.J., & Martin, B.R. (2001). The Economic Benefits of Publicly Funded Basic Research: A Critical Review. *Research Policy, 30,* 509-532.

Sanz-Menendez, L, Bordons, M., & Zulueta, M.A. (2001). Interdisciplinarity as a multidimensional concept: its measure in three different research areas. *Research Evaluation, 10,* 47-58.

Seglen, P. O., & Aksnes, D. W. (2000). Scientific Productivity and Group Size: A Bibliometric Analysis of Norwegian Microbiological Research. *Scientometrics, 49,* 125-143.

Stankiewicz, R. (1979). The size and age of Swedish academic research groups and their scientific performance. In F.M. Andrews (Ed.), *Scientific productivity. The effectiveness of research groups in six countries* (pp. 191-221). Cambridge: Cambridge University Press.

Stephan, P. E., & Levin, S. G. (1997). The critical importance of careers in collaborative scientific research. *Revue d'economie industrielle, 79,* 45-61.

Stephan, P. E., & Levin, S. G. (2001a). Career stage, benchmarking and collective research. *International Journal of Technology Management, 22,* 676-687.

Stephan, P. E., & Levin, S. G. (2001b). Exceptional contributions to US science by the foreign-born and foreign-educated. *Population Research and Policy Review, 20,* 59-79.

Tang, R.; Thelwall, M. (2003). Disciplinary differences in US academic departmental web site interlinking. *Library & Information Science Research, 25,* 437-458.

Thelwall, M. (2002). An initial exploration of the link relationship between UK university websites. *ASLIB Proceedings, 54*(2), 118-126.

Thelwall, M. (2002b). Conceptualizing documentation on the Web: an evaluation of different heuristic-based models for counting links between university web sites. *Journal of the American Society for Information Science and Technology, 53,* 995-1005.

Thelwall, M., & Harries, G. (2004). Do better scholars' web publications have significantly higher online impact? *Journal of the American Society for Information Science and Technology, 55*(2), 149-159.

Thelwall, M., Tang, R., & Price, E. (2003). Linguistic patterns of academic Web use in Western Europe. *Scientometrics, 56,* 417-432.

Thelwall, M., Vaughan, L., & Björneborn, L. (2005, to appear). Webometrics. *Annual Review of Information Science and Technology,* 39.

Thorsteinsdóttir, O. H. (2000). External Research Collaboration in Two Small Science Systems. *Scientometrics, 49,* 145-160.

von Tunzelmann, N., Ranga, M., Martin, B., & Geuna, A. (2003). *The Effects of Size on Research Performance: A SPRU Review.* Retrieved November 22, 2004, from: http://www.sussex.ac.uk/spru/documents/geunaost.pdf

Walsh, J. P., Kucker, S., Maloney, N., & Gabbay, S. (2000). Connecting minds: CMC and scientific work. *Journal of the American Society for Information Science, 51,* 1295-1305.

Weller, A. C. (1996). The Human Genome Project. In S. Y. Crawford, J. M. Hurd, & A. C. Weller (Eds.), *From print to electronic: the transformation of scientific communication* (pp. 35-63). Medford NJ: Information Today Inc.

Whitley, R. (2000). *The intellectual and social organization of the sciences.* 2nd edition, Oxford: University Press.

Widhalm, C., Topolnik, M., Kopcsa, A., Schiebel, E., & Weber, M. (2001). Evaluating patterns of co-operation: application of a bibliometric visualisation tool to the Fourth Framework Programme and the Transport Research Programme. *Research Evaluation, 10,* 129-140.

Williams, F., & Gibson, D. V. (1990). *Technology transfer: a communication perspective.* Newbury Park, London, & New Delhi: Sage.

Wray, K.B. (2003). Is science really a young man's game? *Social Studies of Science, 33,* 137-149.

Younglove-Webb, J., Gray, B., Abdalla, C.W., & Thurow, A.P. (1999). The Dynamics of Multidisciplinary Research Teams in Academia. *The Review of Higher Education, 22,* 425-440.

Ziman, J. (1994). *Prometheus bound. Science in a dynamic steady state.* Cambridge: Cambridge University Press.

Zucker, L. G., Darby, M. R., & Brewer, M. B. (1998). Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises. *American Economic Review*, *88*, 291-316.

Annex

Annex A-1: Academic disciplines based on ISCED 1997

Life sciences
42.1 Biology
42.2 Botany
42.3 Bacteriology
42.4 Toxicology
42.5 Microbiology
42.6 Zoology
42.7 Entomology
42.8 Ornithology
42.9 Genetics
42.10 Biochemistry
42.11 Biophysics
42.12 Other allied life sciences (excluding clinical and veterinary sciences)
Physical Sciences
44.1 Chemistry
44.2 Physics
44.3 Physical geography and other geosciences
44.4 Meteorology and other atmospheric sciences including climatic research
44.5 Marine science
44.6 Palaeoecology
44.7 Other physical sciences
Agriculture, forestry and fishery
62.1 Agriculture
62.2 Crop and livestock production
62.3 Agronomy
62.4 Animal husbandry
62.5 Horticulture and gardening
62.6 Forestry and forest product techniques
62.7 Natural parks
62.8 Wildlife
62.9 Fisheries, fishery science and technology
Veterinary
64.1 Veterinary medicine
64.2 Veterinary assisting
Health (Medicine, medical services, nursing, dental services
72.1 Anatomy
72.2 Epidemiology
72.3 Cytology
72.4 Physiology
72.5 Immunology and immunoaematology
72.6 Pathology
72.7 Anaesthesiology
72.8 Paediatrics

72.9 Obstetrics and gynaecology
72.10 Internal medicine
72.11 Surgery
72.12 Neurology
72.13 Psychiatry
72.14 Radiology
72.15 Ophthalmology
72.16 Public health services
72.17 Hygiene
72.18 Pharmacy
72.19 Pharmacology
72.20 Therapeutics
72.21 Rehabilitation
72.22 Prosthetics
72.23 Optometry
72.24 Nutrition
72.25 Nursing
72.26 Dental services
14 Teacher training and education science
21 Arts
22 Humanities
31 Social and behavioural science
32 Journalism and information
34 Business and administration
38 Law
46 Mathematics and statistics
48 Computing
52 Engineering and engineering trades
54 Manufacturing and processing
58 Architecture and building
76 Social services
81 Personal services
84 Transport services
85.1 Environmental conservation, control and protection
85.2 Air and water pollution control
86 Security services

Source: Own compilation.

Annex A-2: Respondents in the exploratory interviews in alphabetical order

Name	Affiliation(s)	Countries	Duration and date
Professor John Darling	Cancer Research Group and Head of Research Institute in Healthcare Sciences (RIHS), University of Wolverhampton	United Kingdom	40 minutes on April 29, 2005
Professor Jean-Marc Egli	Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Université Louis Pasteur Strasbourg	France	60 minutes on April 28, 2005
Prof. Dr. Ulf-Ingo Flügge	Lehrstuhl Botanik II, Universität zu Köln	Germany	50 minutes on Mai 02, 2005
Dr. Hinrich Gronemeyer	Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Université Louis Pasteur Strasbourg	France	90 minutes on April 28, 2005
Dr. Manfred Heinlein	Institut de Biologie Moléculaire des Plantes, Université Louis Pasteur Strasbourg	France	75 minutes on April 28, 2005
Professor Trevor Hocking	Head of crop and soil sciences research group, University of Wolverhampton	United Kingdom	60 minutes on April 29, 2005
Prof. Alex Kacelnik	Head of the Ecological Behaviour Research Group, Zoology Department, Oxford University	United Kingdom	120 minutes
Prof. Dr. Diethard Tautz	Institute for Genetics, Universität zu Köln	Germany	50 minutes on Mai 02, 2005
Dr. Daniele Werck	Institut de Biologie Moléculaire des Plantes, Université Louis Pasteur Strasbourg	France	75 minutes on April 28, 2005

Source: NetreAct.

Annex A-3: Interview guideline

Notes in advance to the interview – Interviewer: Please read these!

1. Our target group in the survey (and therefore also in the interviews) are the leaders of research teams. If the interviewee has further functions (such as dean of a department) the focus should nevertheless be on the research group.
2. We have to collect the information for the year 2003. If the respondents state that this is very difficult/impossible, we should take a note (and in the worst case ask the questions for 2004). If the evidence that they can't answer the questions for 2003 is overwhelming, we might have to negotiate with IPTS whether 2004 can be chosen instead.
3. The interviews should also help us to find out which questions are unclear, cannot be answered easily or at all, etc. Interviewer, please answer the following questions after the interview yourself:
    a. Which terms were unclear and would need a definition/explanation?
    b. Where can categories be introduced to speed response? Where should lists of categories be extended?
    c. Which questions took a long time to be answered?
    d. Which questions triggered questions on the side of the respondent or needed further explanations to be understood correctly? (note this in the process of the interview)
    e. Which questions were answered unwillingly/ are sensitive?
    f. Where did the respondent try to meet the expectations of the interviewer when answering ("social desirability")?
    g. Was somebody else present/influencing the answers?
    h. Has the interviewee mentioned further important issues or gaps which are not covered in the interview schedule?
4. Please also note the duration of the interview!

<table>
<tr><th>Questions</th><th>Notes</th></tr>
<tr><td>The NetReact project has been commissioned by the Joint Research Centre of the European Union in Seville to explore the impact of the composition of European life science research teams and research collaboration on research output. We are supposed to investigate in particular the role of doctoral students and postdoctoral researchers. Your unit has been selected for first stage exploration, where we are requesting both information on your research team and your assistance in focussing our approach on the key issues.

In line with the research objective, the following questions aim to capture key characteristics of a life sciences research team which you know well. We would like you to answer in respect of the most important unit or group of researchers you have or had responsibility for and which existed at the beginning of 2003.

We define a research team as a group of people who work in a common research area and location. They are part of a larger organisation (university, department, school etc.) and they are recognised from the outside as a separate entity.

Interviewer: “Same location” is not necessarily the same building, but nevertheless they should be “local” (same campus or city) in order to separate it from collaboration.
Ensure a unit is selected and note down any other groups mentioned. If it is inappropriate to select a group which the respondent is responsible for, note the reason. Questions in the present tense can be adapted to refer to the beginning of 2003 if necessary.</td><td></td></tr>
<tr><td>Research team</td><td></td></tr>
<tr><td>1. In what year was the research team founded?</td><td></td></tr>
<tr><td>2. Does your unit contribute to PhD programmes/the education of PhD students?

3. In what disciplines does your research team contribute to PhD education?

Interviewer: Please show the separate list of disciplines from ISCED 1997.</td><td></td></tr>
</table>

4. **Staff of the unit:** What are the approximate numbers at the beginning of 2003 - including yourself - of: a) professors, associate professors, assistant professors b) other faculty with a permanent contract c) post-doctoral researchers with a limited contract d) PhD students e) Masters students f) technicians and administrative staff g) guest researchers and visitors h) others Please include team members which were temporarily on leave and exclude short term guests and visitors (staying less than twelve months).	
5. **Development of the unit:** How has the unit changed within the previous five years (or since its founding)? Has the overall number of staff members grown a lot, grown a little bit, remained constant, declined a little bit, declined a lot? 6. In the case of growth or decline, who contributed to this development in particular? (Interviewer: possible answers are "- -" strong decline, "-" decline, "o" no change, "+" increase, "++" strong increase") • *Permanent scientific staff* • *Post-docs* • *PhD students* • *others (who?)*	
7. Are there any policies that are followed regarding the composition of research teams, e.g. gender-related, presence of scientific disciplines, certain research paradigms?	
8. **Leadership and decision-making:** Please describe the practice of leadership and decision-making in your unit. Who makes the decisions, e.g. on research/teaching duties, research topics, long-term goals, publishing, and who has significant influence on them? The question should lead to a closed question in the survey that assesses leadership style (participative/decentralised – authoritarian/centralised).	

Sources of funding: Please give us some information on the overall funding structure of your research team in 2003. 9. What percentage of your budget was covered through long term core funding (e.g. from your university or other funding bodies)? 10. To what extent did this core funding cover the expenses for • *PhD students* • *Post-docs* Interviewer: ask for 0-25%, 25-50%, 50-75%, 75-100% for each group, PhD students and post-docs 11. What percentage of your budget was covered through further funding from the following sources? • *national research funding agencies* • *EU or other international funding* • *foundations, charities* • *national or regional government and administration* • *companies* • *others*	
Individual information on the head of the unit	
The following part of the interview contains several questions on you/the head of your research team.	
12. What was your position in the research team at the beginning of 2003?	
13. What main responsibilities does this position entail?	
14. When did you take up this position in the unit? (MM:JJJJ)	
Interviewer: Only to heads of research teams 15. When did you obtain the leadership of a research team **for the first time** (MM:JJJJ)?	
Interviewer: **Questions 16-22 can be skipped**, if the interview is subject to time pressure (we need this factual information in the survey, but not in the interviews).	
16. In what discipline(s) do you have a PhD? Interviewer: A separate list of disciplines from ISCED 1997 is provided.	

17. In what country did you obtain your PhD? (In the questionnaire a list of country groups will be provided.)	
18. Have you worked in another country than the current one for 12 months or longer in the course of your scientific career? 19. If yes, in which country/countries?	
20. Within the past five years have you … • *won any scientific awards?* • *served on a major professional committee?* • *served on the editorial board of a scientific journal?* • *organised an international conference?* • *served on a national or international advisory committee?*	
21. Year of birth? 22. Country of birth?	
Doctoral students	
The following part of the interview contains several questions on the doctoral students in your research team and their functions.	
23. **Recruitment of PhD students:** How - that is, by what process – are **applicants sought** for open PhD positions? (e.g. recruiting your own graduates, job ads, asking colleagues)	
24. How are **decisions taken** on recruiting new PhD students? (e.g. committee, team head decides, department head decides) 25. Which of the following can exert significant influence on these recruitment decisions (yes - no): • *head of unit* • *other scientists in unit* • *other university posts or bodies* • *extra-mural bodies?* • *other*	

26. What factors related to applicants are particularly important in the recruitment process of PhD students? Please, let the respondent first provide his/her idea of the important factors and then ask for the roles of the following factors, if they were not mentioned so far: - grades - specific knowledge - discipline of the degree - reputation of the university/department - applicant was recommended - previous working relationship with the applicant's unit	
27. **Functions:** Please tell us, what functions and tasks the PhD students who study and work with your research team have. This question should be included in the questionnaire as a closed Likert scale question. If not mentioned by the respondent, provide the following stimuli: - writing their PhD thesis - learning the practice of scientific research - teaching undergraduates - providing research assistance 28. Do you know of any practice of "abusing" PhD students, e.g. using them as cheap labour without considering their career advancement? If yes, could you describe these practices?	
29. We must collect information about the origin of the most recently recruited 5 PhD students and 5 post docs **active in the unit in the 2001 - 2003 period**, including age, country of birth and qualification. Do you personally have the necessary information to fill in this table? If not, who would it be in this case? Interviewer: Show the table. It does not have to be filled in, we only want to find out, who could provide the information. 30. Would (this person) probably have all the information to hand or have to research this?	

31. **Further career of the PhDs:** How many PhD students graduated in the unit in the last five years/since its creation? 32. Where did your latest five PhD students go after finishing their theses? Try to avoid a bias and remember the successful as well as the not so successful students. This question should be included in the questionnaire as a closed question. If not mentioned by the respondent, provide the following stimuli: - recruited by our team or another team from our parent organisation - post-docs in academia or other public research - private firms - government, international organisations or other - unemployed - don't know	
Postdoctoral researchers	
The following part of the interview contains several questions on the postdoctoral researchers in your research team and their functions.	
33. Is the term "postdoctoral researcher" (post-doc) used in your country/discipline/university and what people are commonly called post-docs? Interviewer: We want to find a possible definition for post-docs for the survey. Please, note the criteria that the respondent mentions for separating post-docs from other researchers. If the respondent asks for our understanding of the term you can provide the following, preliminary definition: We consider as postdoctoral researchers those researchers who have obtained a PhD but not yet tenure or a permanent employment. Usually these are younger scientists who have completed their PhD not too long (less than five years) ago.	
34. **Recruitment of post-docs:** How - that is, by what process – are applicants sought for open post doc positions? (e.g. recruiting your own graduates, job ads, asking colleagues)	

35. How are decisions taken on recruiting new post-docs? (e.g. committee, team head decides, department head decides) 36. Which of the following can exert significant influence on these recruitment decisions (yes - no): • *head of unit* • *other scientists in unit* • *other university posts or bodies* • *extra-mural bodies?* • *other*	
37. What factors related to applicants are particularly important in the recruitment process of post-docs? Please, let the respondent first provide his/her idea of the important factors and then ask for the roles of the following factors, if they were not mentioned so far: - publications - teaching experience - grades - specific knowledge - discipline of the degree - reputation of the university/department - applicant was recommended - previous working relationship with the applicant's unit or the applicant	
38. **Functions:** Please tell us, what functions and tasks the postdoctoral researchers who work in your research team have. This question should be included in the questionnaire as a closed Likert scale question. If not mentioned by the respondent, provide the following stimuli: - improving their knowledge of scientific research - improving their future career opportunities through meaningful research - teaching undergraduate or graduate courses - doing the experimental work of the lab - preparing the publications of the team 39. Do you know of any practice of "abusing" post-docs, e.g. using them as cheap labour without considering their career advancement? If yes, could you describe these practices?	

40. **Further career of the post-docs:** How many post-docs stayed at your unit in the last five years/since its creation? 41. Where did your latest five postdoctoral researchers go after leaving the research team? Try to avoid a bias and remember the successful as well as the not so successful researchers. This question should be included in the questionnaire as a closed question. If not mentioned by the respondent, provide the following stimuli: - another post-doc position in academia or other public research - professor or permanent faculty member - private firm - government, international organisations or other - unemployed - don't know	
Permanent scientists	
In order to reflect the structure of the research team entirely we also would like to ask some questions on the faculty/permanent scientific staff of the unit (excluding yourself and the other types of personnel mentioned so far). You have spoken of xxx faculty members/permanent researchers. Interviewer: Please, insert the sum of the answer to question 4a/b.	
42. How many of the faculty members are female? 43. How many have extended (> 12 months) work/research experience abroad? 44. How many have a PhD/research degree in other disciplines than the main discipline of your unit (e.g. the one in which you contribute to a PhD programme)?	
Research collaboration	
We measure research collaboration through assessing when scientists from different research teams co-author scientific publications.	
45. With how many other research teams did your unit collaborate in 2003? 46. How many of these collaborations would you rate as … - interdisciplinary (your collaborators mainly have a different disciplinary background than your team)?	

- within your main discipline? 47. How has the collaboration activity of the unit changed within the previous five years (or since its founding, if less than five years old)? Has it grown a lot, grown a little bit, remained constant, declined a little bit, declined a lot?	
48. What were the main motivations for the research collaborations of your unit in the year 2003? This question will be included in the questionnaire with a closed question (Likert scale).	
49. Are there any policies that are followed in regard to research collaborations, e.g. related to partners from certain organisations, academia/industry, national/international, etc.?	
Research productivity	
50. What do you think are the most important factors affecting the academic success / publications produced in your unit? Interviewer: If not mentioned by the respondent, ask specifically for the following issues: - Does the composition of a research team play a role? - What characteristics? - Does academic origin of team members have an impact? - Interdisciplinarity? - Heterogeneity/homogeneity in what regards is deemed desirable?	
Role of the Website	
51. What role does your web site play in your research group? 52. What is the purpose of the web site, or what do you hope it to achieve? In particular, does it play a role in attracting PhD students or post-docs?	
Scholarly associations	
53. In order to increase the response rate, a supportive e-mail from a scholarly association in the life sciences might be helpful. Would you know of associations at European/national level and of officers in these associations which we might contact in order to receive this support?	

Annex A-4: Tables with demographic information on PhD students

PhD Students (graduates working at the institute and preparing a PhD)

	Name / alias	Age	Country of origin	Sex	Discipline of Master	University of Master	Year of Master
1							
2							
3							
4							
5							