Conceptualizing and measuring excellence in the social sciences and humanities

By Prof. Peter Weingart and Dr. Holger Schwechheimer,

Institute for Science and Technology Studies – IWT, University of Bielefeld

E-mail: weingart@uni-bielefeld.de

May 2007

This paper was prepared within a Special Support Action "Research Collaboration in the Social and Human Sciences between Europe, Russia,other CIS countries and China" – Global SSH. The Action is supported by the European Community under the Sixth Framework Programme for research, technological development and demonstration (FP6) Thematic Area 'Citizens and governance in a Knowledge-based society' Call FP6-2004-CITIZENS-6 --8.3.3 - Promoting international research and policy cooperation in social sciences and humanities, contract nr. 028997. Additional support to the Action is provided by INTAS under INTAS Thematic Call in Social and Human Sciences 2005, contract nr. 05-1000006-8446. More information on the Action is at www.globalsocialscience.org.

1. The rhetoric of excellence, accountability and evaluation in science policy

Until little more than a decade ago if someone wanted to know who were the best scientists in a certain field he or she would turn to any scientist close by and ask. The judgment of good standing in the respective community of scholars would probably have replied with specifications of the expertise in question. Like scholars institutions' reputations were known to insiders and communicated accordingly. Since then a dramatic change of rhetoric has taken place that is by no means limited to the science policy arena. The rhetoric is part of a much more widespread movement that has been termed the 'auditing society' (Power 1997). The origins of the rhetoric are not accidentally in management. The submission of management to reflexive auditing processes has given rise to similar exercises in virtually all other areas of economic and public decision-making. The language of management (almost exclusively Anglo-Saxon regardless of national and cultural contexts) has penetrated into the public sector and consequently into the realm of science policy. Here, in the funding schemes of research councils and in the administration of universities, it unfolds in the guise of the so-called 'new public management'.

Some of the central tenets of new public management (NPM) are accountability, transparency, efficiency and competition. Correctly implemented these principles of good management are supposed to make the organizations in question, mostly universities, fit for competition. The crucial change in the environment of academic institutions that supported the spread of NPM was, of course, the end of the Cold War, the opening up of the formerly communist countries in Eastern Europe and of China, and the shift to globalized markets, also in higher education. In order to achieve global competition among institutions of higher education and research, benchmarking standards had to be introduced, i.e. standards which would allow for comparisons between these institutions on a global scale. Although it is highly problematic to reduce a complex organization like a university with its heterogeneous structure and multiple functions to a single one-dimensional standard, this has become common practice. The currency of global comparisons and, thus, competition is 'excellence'. Excellence as a concept which is now ubiquitous in science policy is supposed to draw universities into competition with one another, thus creating a stratified landscape of higher education institutions. This new concept is to replace the former system which was oriented mostly to their respective regions or national contexts and whose stratification was largely implicit.

How can excellence in science be determined? Excellence, if it is to achieve its function as a distinguishing criterion, and not to remain an entirely vacuous term, requires some kind of evaluation and assessment. Indeed, a core element of NPM is that organizations are enabled to become evaluated. Evaluation, in turn, must be based on measurable and comparable properties.

This is, as will be argued below, a major departure from the practice of science and its institutions. However, different disciplines are reacting differently to the demands of evaluation which suggests that disciplinary cultures matter. In the following, this essay will explore, first, the significance of the shift to evaluations and the quest for excellence with respect to the institutional fabric of science. Secondly, it will look at the specificity of the social sciences and humanities vis à vis the requirements of measurement. Third, the essay will focus on some tools of measuring excellence in science and their limitations, especially with respect to the social sciences and humanities. Finally, it will deal with the construction of indicators to determine quality of research, and excellence in particular, as well as their pitfalls.

2. From internal attribution of reputation to external determination of excellence

A central object of the sociology of science is the internal mechanism by which the competition among scientists for novelty is combined with the attribution of reputation (Merton 1957; Luhmann 1968). The functionality of this mechanism which is at the core of the institutional autonomy of science is obvious: Since science is a specialized set of languages, theories and methods with which equally specialized new knowledge is produced, all of which are not readily accessible to a lay public, only people with a similar command of knowledge and expertise are able to judge specialized knowledge claims and to assess the competence of their authors. This mechanism is formalized in the institution of 'peer review', i.e. the evaluation of publications and, more recently, research proposals through competent colleagues. Peer review is mostly organized as a qualitative assessment. Judgments of knowledge claims, i.e. of the quality of the research in question, and by implication if not explicitly of the author him/herself, are mostly communicated within the community of specialists. Peer review is based on the assumption that only the 'peers' within that community are competent to judge and evaluate the quality of a particular research.

Peer review is not only the mechanism by which quality of research is being judged, it is at the same time the mechanism by which reputation is attributed. The social structure of science, i.e. the ranking of scientists from a small top of highly reputed ones to the broad bottom of nameless ones, which also translates into control of research resources and differential incomes, is based on the continuous process of peer review that involves evaluations of research proposals and of publications in scholarly journals. It is important here to keep in mind that the process is a going concern. Evaluations are repeated frequently and may lead to varying results. Thus, at least in principle the social structure, i.e. the reputational order in any given community of specialists, is in constant flux.

It is evident that the mechanism of peer review is essentially closed to the outside public because of the specialist nature of the knowledge communicated within the various scientific communities. The only access outsiders have to the internal judgments is by observing the reputation attributed to a scientist. Reputation may be taken as an 'indicator' of quality because it can be assumed that it rests on the collective judgment of competent peers (Luhmann 1970). It has often been argued that the attribution of reputation in science is not entirely based on merit but that it is inescapably cumulative. Robert K. Merton has named the cumulation of reputation, i.e. the tendency to become independent of actual achievement, the 'Matthew effect' (Merton 1973). Systematic analyses have been undertaken to prove empirically the operation of the Matthew effect (Cole & Cole 1973). Even though the results were not entirely conclusive, there can be little doubt that the phenomenon does exist. However, it would be far too radical to conclude that science in its entirety is characterized by the operation of the Matthew effect. If that were the case, it would have ceased to be innovative a long time ago. Rather, it has to be supposed that there is a fragile balance between a conservative social structure maintained by the accumulated reputation – and thus authority – of established scientists and a constant influx of new ideas that are recognized and become the basis for a restructuring of the reputational structure.

While we are not concerned here with the Matthew effect as such, the supposed malfunctioning of the peer review system has been an important reason for the suspicion that science is dominated by 'old boys networks', or, as John Conlan, a representative from Arizona already complained three decades ago, calling the peer review system of the National Science Foundation "an incestuous 'buddy system' that frequently stifles new ideas and scientific breakthroughs, while carving up the multimillion dollar federal research and education pie in a monopoly of grantsmanship" (Subcommittee 1976, 40; c. Cole 1992, 140). The concern especially among policy makers that the evaluation processes within science were beyond their understanding and control have since then grown even more.[1] At the end of the 1980s and in the early 1990s this concern was reflected in a new paradigm of science policy. After the end of the cold war the ideological support for 'free science', i.e. uncontrolled funding for basic research, disappeared. All major Western industrial countries shifted to a régime of 'public accountability' and new public management.

[1] This has been supported to some extent by critique of the system from the scientific community itself which focuses on the alleged conservative function of the peer review mechanism. For an overview of that discussion cf. Weingart 2001, ch.5.

As far as research policy was concerned, this had already been prepared for some years by another dramatic development: Science-based industry closed or downsized their research laboratories and looked to the universities for research co-operations, thus putting pressure on governments and university administrations to make them more receptive to industry's interests and needs. The previous paradigm of science and innovation policy had been based on the so-called cascade model, stipulating that innovative ideas were developed by basic research and were then transferred to application-oriented research, and finally to applied research and development. This model justified the largely uncontrolled operation of academic research, made possible by institutional grants and the virtually autonomous functioning of the peer review system. Subsequently, the cascade model was replaced by more complex schemes of innovation policy that were no longer based on the linear conception but on conceptions of iterative processes between basic and applied research. These suggest an abandonment of the dichotomy altogether (Stokes 1997).

With this conceptual shift the problem of innovation was moved into the focus of science policy more explicitly than before, expressed not least in the emergence of theories of 'national innovation system' (CIS; OECD 1997; 2002). However, if the process of innovation was to be accelerated and directed by state intervention rather than being left to the long term and more or less accidental self-directed processes of basic research, a crucial condition had to be met: The internal mechanisms guiding the process of knowledge production had to be made explicit and accessible to the outside, i.e. to policy makers, research managers, and administrators of funding agencies. Distrusting the internal judgments of researchers expressed in the reputational order and forced to arrive at more specific assessments, the process of evaluation of research, mostly implicit in peer review, had to be translated into quantifiable measures. In other words, the hitherto 'internal', collective, qualitative, and continuous process of evaluation in which not only the quality of certain research results is determined but also the selection of research topics according to their 'internal', i.e. disciplinary relevance is 'externalized'. External 'stakeholders' of research who cannot possibly have the same competence as the peers in respective research fields, are nonetheless enabled to make quality judgments on past research and priority decisions on future research topics if they can base these decisions on proxy indicators.

Obviously, the complexity of the qualitative contents of peer review judgments when translated into indicators has to be radically reduced and simplified. This translation process, consequently, requires knowledge, not to say theories, of the representational relation between the qualitative judgment and the proxy indicator. With regard to science, the most frequently discussed case is the relationship between the quality of research as determined (internally) by the collective judgment of the respective community of peers and the indicator used to indicate quality

(externally), namely, the number of citations of published articles. It is consensus among specialists of bibliometrics that citations represent visibility. It is an additional step to assume that visibility is correlated with quality. An article may be cited and therefore visible because its topic is highly fashionable, because its content is provocative or scandalously wrong, because its author is famous and being cited conveys his/her authority in the cited article etc. None of these reasons for citation are necessarily linked to quality of research (Nicolaisen 2007). This apparent mismatch between the verdict of 'quality' in research (not to mention 'excellence') and the indicator 'number of citations' would be of no importance if one would avoid the term 'quality' and use instead 'visibility'. However, the need for external assessments to legitimate decisions of resource allocations points in the other direction.

The perceived urgency of political decisions leaves little room for methodological reflection and caution. The ubiquitous spread of the term 'excellence' from management into science policy and academia obscures its vacuousness. It has to be given concreteness by measures such as number of citations, prizes obtained, funds acquired etc. All of these are inevitably based on past peer judgments. But cast in numbers, their qualifications and dynamic nature are suppressed and an impression of exactness is conveyed that allows rankings based on two decimal points.

The inherent weakness of indicators of representing only part of the reality that is supposed to be measured points to the goals and purposes of their application. The more consequential the decisions that are based on indicators, the more important it is to consider their methodological flaws. More important yet: Decisions that affect career chances and income of individuals and that are based on indicators which can be influenced by the individuals concerned lead to their manipulation. Economists have analyzed this phenomenon as 'goal displacement' (Frey 2006). Thus, indicators are used most appropriately to *support* qualitative reviews and to *control* their systematic bias if there is one ("incestuous buddy system"). This restrained application of indicators in combination with peer review has been termed 'informed peer review' (Weingart 2005). The use of indicators *in conjunction with peer review* avoids many of the pitfalls of both approaches. If both methods are in agreement, that adds to the validity of the evaluation results. When assessments from both sides differ substantially or contradict each other final judgment has to be based on thorough investigation.

Even if one disregards the fact that the very term 'excellence' comes from the managerial world and is, as such, meaningless, the 'externalization' of evaluation in science means that comparisons have to be established: between individuals and institutions. The focus shifts to quantitative comparisons, away from judgments of research proper. In lack of a substantive definition of excellence, it has to be understood as a relational term, and indeed, the recent use of the term in science and higher education policy reveals its actual function. It is not to

determine what kind of research or who among any group of scholars is 'excellent'. Rather, the objective is to introduce competition among researchers and institutions, and, by sanctioning certain products, to create a hierarchy among them. Excellence has to be scarce. At the top is only room for a few, the steeper and narrower the hierarchy the better.

3. The specificities of the social sciences and humanities vis à vis the natural sciences

The aim of this analysis is not to establish hierarchies and to initiate competition but rather, first of all, to identify scholars and institutions of the social sciences and humanities in countries of the Commonwealth of Independent States (CIS), i.e. countries of the former Soviet Union as well as in China. That task is difficult enough because the respective communities have been isolated from communication with their respective counterparts in the West. Then, after having identified individuals and institutions, the next task is to assess them, in order to, thirdly, mobilize, i.e. integrate them into the international community of scholars. To accomplish these tasks it is of foremost importance to point to some specific characteristics of the social sciences and humanities that set them aside from the natural sciences. Since discussions about the differences between the different disciplines have been an on-going concern since the 19th century, only a few relevant points shall be mentioned as a reminder.

Most pertinent for the task at hand is the *parochial* nature of the social sciences and humanities. This is not meant pejoratively. In a recent study of international social science research commissioned by the UK's ESRC the authors state: "Most social science research is done within national and local boundaries, and, most often, by individual scholars rather than the research teams which populate medical and natural science research" (Forbes/Abrams 2004). The obvious reason is in the nature of subject matters. The laws governing matter are universal and can be studied anywhere. The subjects dealt with by the social sciences are much more constrained: the social structure of a given society, the impact of a tax reduction on political opinions of a given electorate etc. This is probably even more true for the humanities, many of which – such as the philology – are national by definition.

The different nature of the subject matters between the natural sciences and SS&H accounts for one highly important characteristic: the cumulativeness of the natural sciences. SS&H are *non-cumulative* which means that creativity, progress and excellence are much harder to determine. Judgments about the value of specific contributions and the creativity of their authors are often contradictory, varying along the lines of schools of thought. Cumulativeness is attributed to the degree disciplines are paradigmatic. Disciplines that are paradigmatic in the Kuhnian sense have a higher consensus over the relevance of research problems and the evaluation of answers. Typically, the principal publication is the journal article rather than the monograph as in the

SS&H (although there are also great differences between different fields). This means that the natural sciences grow faster and the rate of obsolescence of their knowledge is faster, too. 'Schools of thought' are virtually absent (Whitley 2000).

For all these reasons the SS&H are not in the purview of a utilitarian science policy. Their products and immediate impact can rarely be measured, be it in terms of demonstrable effects such as in the political arena, be it in monetary terms, i.e. as in the economic arena contributing to the development of new technology and to the balance of trade. In times when the test of proven utility has assumed priority in science policy, the SS&H have continued difficulties in mobilizing political support for themselves. Their success cannot be measured in terms of 'return on investment'.

Rather, the SS&H are *reflexive* disciplines. Their subject matters have to do with some aspect of human social action. This means that they are not only in constant flux but they also change under observation, and the criteria of relevance and observation are subject themselves to continuous re-interpretation and assessment. To the extent that they have an impact at all it is usually indirect and long term, thus difficult to attribute to particular 'discoveries' or achievements. What Iverson states for development studies is certainly true for other areas of the SS&H as well: "Interventions within complex systems are embedded in, and affected by, the uniqueness of time and place…Insofar as multiple and often unknown confounding variables are the norm, complex systems present a serious obstacle for attribution" (Iverson 2003, 36; cf. Carden 2004).

A few fields among the SS&H are partially exempt from this description and have more direct instrumental functions: opinion polling, operant conditioning, and epidemiology to name a few examples. But even in these research fields, basic assumptions, categories and theorems are by no means uncontested. They only seem less ambiguous and, thus, more 'relevant' because they are more directly linked to (political) action. Utility does not even seem to depend on the truth value of the knowledge in question: opinion polls may be wrong, and economic forecasts of growth are systematically adjusted after the fact, and yet both instruments are in strong demand from policymakers (Schmid 2005).

Conversely, the indirect and long term effects of ideas emanating in the social sciences or humanities may be dramatic. Economists like Marx, Keynes, Schumpeter and Friedman have, each in their own way, had enormous influence in shaping their respective societies. Public intellectuals like Sartre, Habermas, or Foucault, philosophers like Heidegger or Russel, and psychiatrists like Freud, Jung and Skinner have shaped the thinking of their times at least as much as, if not more than, their natural scientist contemporaries.

Examples such as these reveal immediately that neither utility nor direct measurable impact is a reliable yardstick for assessing the value of the SS&H. Yet, governments and the media require that the sciences prove their value to the public in terms of measurable utility. Thus, the only way for the SS&H to escape this dilemma is to reject utilitarian expectations altogether (a strategy followed by the German Science Council; Wissenschaftsrat 2005).

Patterns of international collaboration

Does all this mean that the SS&H are by their very nature bound to their respective cultures, societies or nation, unable to communicate, let alone collaborate across the respective demarcation lines? Such a picture would be much too simple. International cooperation is not unknown in the SS&H. International scholarly societies such as the International Sociological Association (ISA) under the auspices of UNESCO/ICSU organize congresses and journals. Various subdisciplines and specialties are also organized internationally, and members read, correspond and co-author across the national boundaries. In what ways are the SS&H different from the natural sciences as far as international collaboration is concerned?

Contrary to common belief, international cooperation in the natural sciences is by no means distributed equally among nations but shows uneven distribution between North and South, East and West (Engels et al. 2005 and the literature cited therein). The same is true for the SS&H. A study of collaboration patterns in the SS&H – the only one so far available to our knowledge – with a focus on Canada reveals interesting patterns. First of all, "International collaboration is growing steadily in the natural sciences and engineering (NSE) and social sciences but remains unchanged in the humanities….The vast majority of articles in the NSE (90%) are written by more than one author. In the social sciences, the proportion of multi-author articles more than doubled during the period 1980-2002, rising from 30% to nearly 70%, but in the humanities the proportion remained stable at about 10%. While these figures indicate three distinct trends, they also suggest that the collaborative practices of scholars in the social sciences correspond more closely to those in the natural sciences than to those in the humanities. (Larivière et al. 2004, IV)". The differences between the natural sciences and the SS&H exist also among the the SS&H, as this study shows. "Overall, psychology and economics and administration were the disciplines with the strongest collaboration, followed by social sciences, education, and then law. In the humanities, history was the discipline in which collaborative activities were most frequent, but the rate was still low. In the humanities and literature, collaboration was a marginal phenomenon. Note that, overall, the disciplines with the highest collaboration rates are also the ones in which journal articles are the main medium of knowledge dissemination" (ibid.). Not surprisingly, the study also hints at reasons: "The bonds of language and history influence the

choice of international partners" which is "illustrated by the special ties between Quebec and France and the higher percentages for the U.K., Australia and France in history and literature" (ibid.). Looking at inter-institutional collaboration the same pattern emerges again. "Physical distance as well as language does play a structuring role in the choice of collaborators" (Larivière et al., 2006, 531).

Even though the focus of the study is Canada, there is no reason to assume that the results would differ for other countries. They corroborate common wisdom about the SS&H, namely that they are more culture bound than the natural and engineering sciences (NES). Although the growing international collaboration in the social sciences also shows that they can become more internationalized, the humanities are in a different category altogether. This raises the question if the push for internationalization favored by evaluation schemes is really warranted equally for all fields. But the results of the study have another significance: the bulk of scientific collaboration, also in the SS&H, takes place between the major Western industrial countries. The CIS and Russia appear hardly or not at all in the relevant data. China and South Korea (!) are somewhat more visible. Because of their cultural embedding and their strong ties to language, changes in collaborative patterns are slow.[2] This is a warning against overly optimistic expectations about the effects of quick mobilization and simple measures to initiate it.[3]

Rather, it appears that mobilization efforts directed at countries whose SS&H have not yet appeared on the map of international collaboration will have to look very carefully at specialties within the SS&H and look at thematic needs and opportunities. A recent study of collaboration patterns from the UK perspective, although focused on strengths and weaknesses of social science research, showed that future prospects of international social science were seen to vary between thematic areas (Forbes, Abrams 2004, 239p).

4. The SS&H in the light of bibliometric indicators

[2] This appears to be particularly pertinent in the case of the CIS countries. G. Roll attributes the low international visibility to 1) the historically and politically constituted territorial divide in the reproduction of social knowledge, 2) the high average age and scientific training of CIS researchers that does not motivate international publishing, 3) the language barrier (Comment by G. Roll).

[3] "The institutional systems of science management and evaluation currently existing in the CIS countries do not motivate scientists to publish in internationally recognised journals. In the Southern Caucasus (Armenia, Azerbaijan and Georgia) science management systems are still under development. In the largest country in the region, the Russian Federation, only publications in Russian-language journals are considered in the evaluation of Russian scientists. PhD candidates are required to present a number of scientific publications as a condition of admission to the PhD defence. However, only publications in Russian and in a fixed list of Russian journals are considered. Publications in international journals are disregarded (pers. comm. Ms. Natalia Alexeeva, PhD candidate at St. Petersburg State University, Russia). Russia also has its own Russian-language citation index used for the evaluation of the scientific output" (Comment by G. Roll).

The easiest way to identify prominent researchers, important research results, and institutions fostering good research is by way of bibliometric analysis. For that reason so called bibliometric indicators have been developed as devices to evaluate research. The principal source of information for bibliometric analyses is the *Social Sciences Citation Index* (SSCI), and no other database "even comes close" to this index (Hicks 1999, 193). The same may be said for the humanities for which the *Arts and Humanities Citation Index* (A&HCI) is the respective databank. ISI currently indexes roughly 1800 journals in the social sciences, 1200 in the arts and humanities, compared to 5500 in the sciences. The reason for the superiority of these databanks over their competitors is the unique combination of information about the author(s) of a given article, his/her institutional address, and the citations to articles. This means that searches can be made targeting authors, their institutions and/or number of citations received by an article.
This data bank has also been used as a tool for the evaluation of research as it is reflected in publications and for studies of communication patterns, i.e. of social structures in science generally. For this purpose so-called bibliometric indicators have been constructed.
The most important bibliometric indicators for activity (publications) and impact (citations) are:

P: Number of publications

C: Number of received citations

CPP: Citations per publication

CPP/FCS_m: Normalized citation rate (against Field Citation Score mean)

The most basic indicators are:

- number of publications – indicating the *activity* in formal communication, usually being taken as an indicator of research *productivity;*
- number of citations – indicating the visibility or impact of research but usually being taken as an indicator of the quality of research.

Publication counts are often criticized because of the the broad variation in type, volume and importance of content. This deficit can be partly corrected by excluding problematic document types (e.g. letters), but the problem of activity as a neutral dimension remains. A better measure which probably comes nearer to 'quality' is 'impact' or visibility which can be measured by the number of citations of an author's work. To be comparable over different fields of research or disciplinary profiles of institutions, the average citation rate of the field in question has to be considered. The probability of being cited by colleagues depends not only on the importance and usefulness of the information but also on the culture of communication practice of the respective scientific communities. An average publication in biochemistry will be cited much more often than an outstanding work from a small specialty in mathematics. To normalize the citation rates (citations per publication) the absolute citation count is divided by the average

citation rate of all publications of the same discipline or journal (to get a more narrow comparison value) from the same year of publication. If computed for a sufficient number of publications, this indicator is widely accepted in bibliometrics as a reliable measure for visibility in most areas of the natural sciences. In the social sciences and more so in the humanities, this form of application is highly problematic, because of the bad coverage in the Citation Indices of books which are the most important publication types in many fields in these disciplines.
While bibliometric analyses have become a well established instrument in the evaluation of the natural sciences, "the use of the SSCI for bibliometric applications is covered with obscurity and myths…The main argument against applying bibliometric techniques in the evaluation of social sciences research has always been (and still is) the (poor) coverage …both in terms of literature covered as well as in the scope of the journals…" (van Leeuwen, 2006, 133). Much research has been done on SSCI coverage, its limitations and future prospects, and the main results will be presented here. Unfortunately, the situation is less well known in the humanities.
A decisive factor determining coverage is the type of publishing. Books are not covered by the indices but still have an important place in the publishing activity in the SS&H. A large variety of studies estimating the share of articles compared to books provides a broad range of figures because they are usually national in scope and/or pertain to specific databanks and sub-disciplines. The share of journal articles in the social sciences may be as high as 60% and as low as 40%. They all concur in the result that "journal-based bibliometric indicators will be based on a smaller fraction of research output in the social sciences than in the natural sciences" (Hicks 1999, 195).
One of the few in-depth studies of publication patterns in the humanities showed remarkable differences between disciplines such as English and American studies, German studies and History. If one controls for the growth of the respective communities, looking at the average productivity per professor, the article as the major type of publication gains in History, remains fairly stable in English and American studies and loses ground in German studies in the period between 1954 and 1984 (Weingart et al. 1991, 283pp). The differences in types of literature between fields also accounts for their coverage in the SSCI and A&HCI. The more 'scientific' the respective discipline aspires to be, the higher the share of the journal article, and, thus, the better ISI coverage. "Economics and psychology literature (is) the best covered" (Hicks, op. cit., 195 and pp. for more elaborate data).
An important issue of concern is the role of books in the SS&H. In the natural sciences the rate of knowledge production can be faster because of the higher degree of consensus. Thus, the danger of being anticipated by another author is real while social scientists rarely have to fear that. More time may be invested in the production of monographs. Although the ISI indices do

not cover books, "producers of social-science indicators are forced to admit that the best social science is often found in books. This should be reflected in citation rates, and empirical studies find that it is" (Hicks, op.cit. 197). Individual books may be very highly cited, and taken altogether they account for about 40% of citations (Ibid., 201). In some fields in the humanities the importance of books may be even greater than in the social sciences as the above mentioned data from the German study suggest. The omission of books from ISI's databanks may not preclude bibliometric analyses of the SS&H altogether, but suggests that an important share of the knowledge production in these fields is not covered.

Another, equally critical issue is the national bias of ISI's indices. Several studies have focused on this problem since the common view is that the indices are biased in favor of US authors. A comparison between SSCI and UNESCO's *World List of Social Science Periodicals* shows that not only is that list much longer than SSCI's, but it also over-represents the US and the UK while France, Germany, and the rest of the world are not well covered (Schoepflin 1990, 181). More recent studies show that this situation has changed somewhat and will probably continue to change. The position of the US in the SSCI is relatively stable over the period 1991 – 2003 while the output of the European countries shows a large increase. The Netherlands, Denmark and Sweden have more than doubled their output, France and Germany have nearly twice as much output by 2003 as compared to 1991. "The share of papers not originating from the US is increasing, thus repelling the dominant position of the US within the SSCI" (van Leeuwen 2006, 141).

These numbers seem to indicate a growing internationalization in the sense of a stronger representation of countries other than the US in the SSCI. They do not necessarily indicate more intense international collaboration among social scientists. Impressionistic data and inferences suggest a cautious internationalization of the social sciences. Funding programs on the EU but also on national levels are nudging scientists to publish in English in high reputation journals leading to a higher share of publications covered by the SSCI. These programs are part of the greater effort to integrate research institutions in a "European Research/ Knowledge Area". In the case of the socio-economic domain, the integration is "still very low compared with other research fields" (Cappellin, 2004, 208).

Van Leeuwen shows for output of geographic regional classes (articles with single author US or cooperating with US vs. no US involved) that in SSCI fields (compared to SCI fields) the US still "covers nearly 60% of the output, either singular or through scientific cooperation" (van Leeuwen 2006, 147). His analysis of citation flows in economics reveals the 'single USA' publications to 'single USA publications' as the largest with 45% decreasing to 35% for the

period 1991-1995. A surprising finding is that the share of 'no USA involvement' publications in SCI fields is larger than in the SSCI fields.

Some of van Leeuwen's findings are particularly relevant for the decision to apply bibliometric methods or not. "The publication cultures within the social sciences differ across disciplines", van Leeuwen notes, while the "results on country level show a strong resemblance of the publication behavior across nations". The differences across disciplines limit the reliability and validity and the macro-bibliometric approach within the domain of the social sciences. In micro-bibliometric analyses, i.e. within a field and within a country, the techniques can be used as long as results are not over-interpreted and are checked by peer review (van Leeuwen 2006, 151). Summarizing the results on applying bibliometric techniques to the SS&H,it can be said that one cannot rely on their reliability and validity in the same way as in the natural sciences because of: the non-paradigmatic nature of most fields in the social sciences and humanities, the heterogeneity of publication behaviours between fields in the SS&H, and the insufficient coverage in the SSCI and A&HCI . The latter is changing, at least for the social sciences, as a result of an increasing internationalization due to incentives for non-English speaking authors to publish in English. This is particularly true for the European countries where funding programs promote publication in English in order to achieve integration of the European research area.

Coverage of SS&H of the 'Newly Independent States' (CIS)

With the end of the Cold War "the transitions of East and Central European nations have allowed their scholars to participate in international science, not only by freeing communication and travel but also by reducing …'political bias' in their work…Pre-transition, the SSCI missed 90% of Polish sociologists; post transition it missed only 30%..." (Hicks, op. cit. 207). This optimistic view may hold for those countries that have entered the EU. However, the member countries of the CIS are more remote from the international communication networks, and they have much less well developed, smaller communities of researcher in SS&H. In order to get an idea of the size of the publication output it is useful to look at the number of publications in all fields for each country in the *Web of Science* (WoS; cf. Tables 1-12). There we see that only the Ukraine and Russia have four and five digit numbers in the period between 1996 and 2005. The remaining countries are in the tens (Turkmenistan, Tajikistan) or hundreds. The next step is to look at the publications of each country published in the SSCI and the A&HCI (SSAHCI). Here it becomes obvious that in all countries except Russia and the Ukraine, the respective numbers are in the teens or single digits. This means, in effect, that one cannot speak of SS&H communities, but most likely of individual scholars who work more or less in isolation. The numbers themselves do not reveal any trend, neither towards higher numbers in the SSCI and

the A&HCI nor the reverse – with the exception of Russia and the Ukraine where the absolute numbers of articles published and included in the two indices show a downward trend. What remains unknown, however, is the actual number of scholars and their output because one cannot control for the percentage of coverage of CIS articles in the SSCI and A&HCI. This would require an independent stocktaking of researchers and their work. Under such circumstances the application of bibliometric techniques is out of the question.

5. Indicators of the quality of research? – Encompassing assessments of fields

The evaluation of the quality of research and, thus, the choice of methods and indicators depends on the (implicit or explicit) concept the evaluator has of the research process. If the quality of research is seen to reside in the creative genius of the individual scientists, the focus is on the search for factors promoting individual creativity. Psychologists have focused on creative acts and on situations in which they occur. Creativity is clearly a psychological topic (cf. Sternberg 1999). However important insights into the cognitive qualities of creativity may be, they are clearly insufficient when it comes to how to explain creative achievements in science between different fields at different times in different cultural contexts. Typically, the reference for the definition of creativity in science are past achievements recognized by the respective community (e.g. Darwin's formulation of evolution, Einstein's relativity theory), and attributed to individual scientists (Heinze et al. 2007). This approach ignores the fact that many so-called discoveries are not recognized as such for a long time – Darwin and Einstein are both cases in point – and that they are rarely – Einstein's relativity theory is an exception – the product of a single mind but have many precursors and sometimes even 'co-discoverers' – as is the case for Darwin's theory of evolution.

Although suffering to some extent from the same methodological problem, sociologists have taken the route of trying to identify organizational conditions that are conducive to creative achievements in research. The Hollingsworths look at particularly successful institutions, universities like the Rockefeller Institute or the California Institute of Technology, that both boast a long series of outstanding research achievements that have gained a plethora of prizes. They avoid the circular definition of 'creativity' to some extent by taking 'major discoveries' as their referent, defined as a "finding or process, generally preceded by numerous 'small' advances, that solved a particular problem and in turn led 'to a number of smaller advances, based on the newly discovered principle'" (Hollingsworth&Hollingsworth, 1999, 216). The 'major discoveries' are identified by prizes, including nominations for the Nobel Prize because of the scarcity of cases if only the awarded prizes were taken. They then look at the organizational features of these institutions that (may) have contributed to establishing such

successful traditions of creative research. The indicators deduced from the analysis are: 1) *scientific diversity:* variety of specialties and people with different disciplinary backgrounds, 2) *depth:* number of scientists in each area of diversity, diversity of talents in each scientific area, 3) *differentiation*: number of departments and other kinds of units, 4) *hierarchical and bureaucratic coordination*: standardization of rules and procedures, 5) *integration of multidisciplinary perspectives:* across specialties the frequency and intensity of interaction, 6) *visionary leadership: capacity for understanding direction in which scientific research is moving and integrating scientific diversity*, 7) *quality*, proportion of scientists in the nation's most prestigious academy of science, research funding per scientist (Hollingsworth& Hollingsworth 1999, 219p). If one accepts the methodology of this approach, it is evident that all indicators point to *diversity in communication facilitated by organizational means* as the principal overall condition of successful research. Although the analysis is focused on biomedical research, it may be safely assumed that this general principle may be extended to the SS&H as well.

The rationale of comparing psychological and sociological approaches to the study of creative research is to demonstrate that the choice of indicators is guided by a theory of scientific research and the emergence of quality standards. Psychological theories assume that 'creativity' is largely situated in the individual; sociological theories assume that 'discoveries' are induced by organizational conditions and communicative networks. Philosophical and historical theories focus on the intellectual traditions and contexts generating new ideas. These theories, very often implicit, determine the range of conditions and by implication of indicators that are considered relevant in bringing about high quality research. However, as the theories move from (mysterious) individual qualities to complex social systems, the range of possible factors becomes limitless, and the theory – if there is one at all – becomes fuzzy.

Current efforts to evaluate specific disciplines, institutions, and researchers initiated by governments in the attempt to cut costs and induce competition by creating hierarchies are most often not guided by any theory. Instead, they are either determined by administrative and political pragmatism, resorting to very few easily, manageable indicators, or they are constructed under the influence of the institutions and researchers to be evaluated. In case of the latter their interests override theoretical considerations leading to unwieldy collections of indicators whose contribution to quality in research is unknown.

The evaluation of the humanities and sociology by the German Science Council - experiences

The German Science Council (*Wissenschaftsrat*) under the mandate to formulate recommendations for policymaking in the German higher education and research system has

undertaken a major review of the humanities (*Geisteswissenschaften*), and is currently carrying out an encompassing pilot evaluation of two disciplines: chemistry and sociology. In both, the report on the humanities and in the evaluation of sociology, qualitative and quantitative elements are combined but with different weights, in line with the different goals of the two activities. The evaluation of the humanities was an effort to take stock of twenty years of development in order to update policy in the area. One part was the collection of data for a general overview; a second part was devoted to the evaluation of six humanities research centers that were established in 1996 to enable the continuation of research groups that had previously belonged to the former East German Academy of Science and had been evaluated positively after reunification. Thus, the Council's report contains the two principal types of evaluation: of an entire field of research on a national level, and of specific research institutions.

It is evident that the national evaluation is strongly influenced by the current discourse on university reform, especially the introduction of competition between universities and the inevitable construction of specific 'profiles', realized mostly by the imposition of quantitative indicators. Two sets of indicators figure prominently: the attraction of external funds and publication measures. Both measures are considered problematic, partly for the reasons already discussed above. It was found that most journals in the humanities (at least in Germany) are not peer reviewed, most likely because they are not considered an important publication outlet in the first place, and for lack of consensual standards.

While the attraction of external funds as such is an inappropriate measure – especially if compared to the natural sciences – because much research in the humanities does not require large amounts of money (notable exceptions are long term text editions and archaeological excavations), the establishment of 'centers' for research and teaching in universities is considered an important structural prerequisite to promote interdisciplinary work in the humanities (Wissenschaftsrat 2005, 46pp).

In its recommendations the Science Council, thus, stresses the importance of thematic foci and centers since it considers 'multidisciplinary environments' and opportunities of interdisciplinary exchange as pre-conditions of successful research in the humanities. In exceptional cases such centers could also be established nationally or at least between cooperating universities. Because of a scarcity of available resources it also suggests that universities which opt to include the humanities among its departments should cover a core of subjects in five dimensions: 1. languages/texts, 2. image/music/theater, 3. history/society, 4. knowledge/ethics/religion, and 5. a representation of non-European areas of knowledge. They can then choose the disciplinary profile. Finally, the council strongly recommends that the respective scholarly associations develop criteria of evaluation that can be applied in the humanities. In this context it stresses the

urgency ofo introducing peer review as a standard mechanism of quality control (Wissenschaftsrat 2005, 49-50).

The (largely qualitative) evaluation of the six centers follows a fixed pattern. The centers are described in terms of their history, objectives and missions, their work program, organization, staff and funding. With respect to the latter two points the director(s) of the centers can comment whether they consider the resources sufficient to attain the stated goals of the institution. Further descriptions concern the publication output, attendance of conferences, visitors, and co-operations with other institutions as well as the promotion of young scholars. The actual evaluation is carried out by a group of experts (an interdisciplinary evaluation team) that visits the institution and discusses the self-report. The critical points are usually the coherence of the research program (or lack of it), the funding situation and its prospects, the research output including the number of completed dissertations, and the integration into the national and international community in the light of co-operations and visitors. Although quantitative data are provided for all aspects they are not formalized into indicators (Wissenschaftsrat 2005, 149pp.).

The Science Council's evaluation of '*sociology'* is a pilot study that is supposed to generate information about the feasibility of certain indicators for future, more formal evaluation processes. Thus, the rationale – and the temptation – is to collect as much data as possible. The council's interest is to obtain a complete picture of the research situation (both descriptive and evaluative) of the discipline. All academic institutions (59 universities and major research institutions) where sociologists work are included in a poll which asks over 120 questions. The questions were formulated by the council staff in collaboration with a team of evaluators. All of the latter are sociologists.

One part of the evaluation seeks to obtain a description of the institutions, a second one an account of research units (260) within institutions in order to achieve some degree of comparability. Underlying this is a 'quality model' which consists of six dimensions: research quality, impact/effectiveness, efficiency, promotion of young researchers, transfer to areas of application, mediation and diffusion of knowledge. In each dimension quantitative and qualitative indicators are represented.

The questions cover the description of the organizational structure and research foci, staff and funding, co-operations and visitors, mechanisms of quality control, external funding, doctoral degrees conferred and 'habilitations', research products and publications in non-academic media, academic and non-academic positions. Each unit submits between 5 and 7 publications for evaluation, depending on its size. Parallel to that all publications produced under the address of the institution will be searched in specific data banks for a publication count. However, no

citation analysis will be undertaken for fear that the coverage is too small. Instead, the evaluation team will assess a selection (decided by the members of the research units) of ca. 860 publications. It is also responsible for rating each research unit on the basis of all indicators. The rationale is that of 'informed peer review'.

The process in which the questionnaire was formulated and decisions were made on the inclusion or exclusion of particular indicators made clear that it is not guided by a specific theory about how science (or sociology) operates and which indicators best reflect that. Rather, the concern was – guided by a variety of implicit views – which indicators should be included so that the discipline (or certain institutions or research units) would appear in a positive light. As a result the questionnaire turned out to be as long as it is. At the time of writing, the results are not yet in, but so much can be said already: the effort is extremely costly, involving many uncounted work hours not only of the evaluation team but also of the staff responsible for the collection of data in the individual institutions. The probability of obtaining false data and/or creating artefacts is high given the amount collected, the varied availability, and the inevitable mistakes which might be made by those who collect them. The degree of contingency of ratings that are finally assigned to the institutions is also considerable given the number of subjective judgments involved. The large volume of information suggests an exactness that cannot possibly be achieved in the actual process of evaluation. Ultimately, the question will have to be if the same results cannot be obtained at far lower costs.

On the positive side it must be said that the evaluation will most probably generate an impressive amount of data on the organization of sociology as a discipline, on the profile of research and on the productivity of sociologists. Even if one can have doubts about the relationship between costs and actual benefits in terms of reliable results, once the data are available, a lot can be learned from them.

The evaluation process of sociology points to a lesson that should guide similar efforts in the future. Any attempt to assess the state, achievements and comparative quality of a research field, notably in SS&H where the applicability of quantitative data is problematic, should attempt to balance the presumed exactitude of information with the effort spent on obtaining the pertinent data. This pragmatic principle applies in particular the larger the unit under observation, when data collection is time consuming and, thus, costly, when the collection of data is often inevitably incomplete and, due to reliance on many diverse sources, also fraught with mistakes.

The impact of political environment and infrastructure – Assessments of the social sciences in Eastern Europe and Asia

In 2003 the German *InformationsZentrum Sozialwissenschaften* (IZ) invited social scientists from 13 Central- and East European countries (CIS and Russia are labelled East European for short) to assess their needs of information and to give feed back on services provided by GESIS (Society of Social Science Infrastructure Institutions, a German government research institution) to East European social sciences. The findings of this meeting appear relevant here because they reflect the self-perception of East European social scientists, of their needs and deficiencies especially with respect to two aspects: political environment and infrastructure. It is obvious that the SS&H cannot flourish in an oppressive or corrupt political environment, and it is equally obvious that a good infrastructure is a precondition to high quality research. The IZ report is a unique source in the sense that it highlights these external conditions as indicators of the state of the SS&H and that it does so by reflecting the self-perception of the Central and East European scholars. To some extent it updates UNESCO's World Social Science Report 1999, although it is much smaller and the OECD's "The Social Sciences at a Turning Point?" of the same year (UNESCO 1999; OECD, 1999).

In some East European countries the legacy of the pre–1989 régime is still alive. That means that the old 'research elite' still plays a leading role, i.e. it controls the access to the existing infrastructure, to funds and international contacts, often preventing or at least hampering innovative research and international cooperation of young researchers. The situation of scholars in Eastern Europe is deemed less constrained than in the former Soviet Union where a lack of English, little knowledge about Western scholarship, and a general scepticism toward the values of Western scholarship have had a retarding effect on the development of the social sciences (UNESCO 1999, 92p.). This is particularly true in a country like Belarus where the freedom of research is still "extremely constrained" (IZ 2004, 5).

Political interference has unfortunate repercussions in systems where funds for the SS&H are scarce. The East European countries spend less than 1% of GDP on research, and the situation in Belarus and the Ukraine is considered more serious than in Slovenia or Poland. The already scarce funds for the SS&H are often controlled by entrenched 'old boys networks' that prevent their distribution according to transparent rules. What the IZ country report states for Romania is considered generalizable for the region: "The general perception is that all state-administered opportunities (grants, internships, etc.) are distributed on the basis of connections and/or bribes. Even if this perception is not 100% correct, it discourages individuals without 'connections' from applying. The results are lack of progress in the reform of the social sciences as well as brain drain" (IZ 2004, 12).

This means that there is a strong dependency on foreign research funds. Here the entry into the EU of some of the countries that were formerly connected with one another in an East-East cooperation, has, according to observers from the region, resulted in a further rupture of these links beyond what has already occurred since 1990. The new member states will profit from their entry while the CIS states will risk being left behind. This may lead to a re-orientation of research in the new member states towards EU framework programs and the abandonment of cross-national East-East comparative studies. The facilitated mobility will also increase the danger of brain drain.

Another grave problem is the highly structured form or complete lack of cooperation both within these countries as well as internationally. Lack of funds and lack of a well organized infrastructure prevent social scientists from obtaining information both about the activities of their colleagues in the same country and about the international social science community. There are no unified and comprehensive national information systems for social science research (except for certain disciplines and institutions).[4] Access to the international literature is very limited due to financial constraints of the libraries and publishing companies. The need for international literature is huge (IZ 2004, 21). Part of the problem is the lack of internet databases for publications.

National and international cooperation is hampered by the very limited availability of communication technology (except email which is used widely).Use of the internet is constrained by lack of funds. This has an effect both on the versatility of using the internet and the visibility that the internet offers to individuals and institutions.[5] A small test can illustrate

[4] "It is important to note that according to one of the PREST developed scenarios of the RTD systems development in the CIS countries (see at http://prest.mbs.ac.uk/prest/scope/CIS_reports.htm), there is a two-tier development of research capacities in the CIS countries. This scenario seems to describe rather well the current state of the RTD system in Russia and other CIS. There are a few strong elite research and higher education institutions, internationally recognised „islands of excellence", which use effectively international funding, have good scientific infrastructures and produce most of the research publications (e.g. Higher School of Economics, New School of Economics, and European University in St. Petersburg as well as a few other institutions in Russia), and have strong scholarly connections with the top universities in the USA. On the other hand, there are a huge number of regional institutions and universities almost totally oriented to the local market and work in Russian or other local languages only. Those institutions face the funding and research infrastructure challenges" (Comment by G. Roll).

[5] G. Roll comments: "The use of communication technologies varies greatly from country to country. In some regions and CIS countries international donors developed strong Internet communication infrastructure. In some countries of Central Asia such as Turkmenistan, Uzbekistan, governments imposed restrictions to access to international Internet sources. In Russia …quite rich information and news are available (on the Internet – PW/HS) however this does not apply to the availability of scientific journals: only 22% of SSH journals in Russia are available also on Internet… The difficulty is connected to the inertia of the scientific journals editing teams… publishing is organised only in hard copy; requirements of references in published papers are minimal. Also due to the fact that the all-CIS and all-Russia scientific communication system is not working well, publications appear in local and regional journals rather in than in national ones."

this. We searched the Internet in order to determine whether it offered a basis to assess the literature output of the CIS in SS&H independent from the Web of Science and thereby reveal deficiencies of the WoS. For two states, Armenia and Ukraine all institutions mentioned in the text of Pipiya/Roll were searched in the internet. For some of the remaining states we took spot checks. The result is the same for all: the presence of the institutions in the internet is very sketchy and remains on a general descriptive level. Detailed information about research projects and publications are often missing completely. For the two countries whose institutions were covered completely, we obtained the following picture: for 15 of the 50 institutions no pages could be found at all. Four of the pages found were neither in English nor in French. Only on eight of the 35 pages identified we found publication lists of the respective institution. All in all, in only 10% of the cases truly informative publication lists were available that could also be used as a basis for evaluation.

Most of the national communication takes place in the existing journals that have survived and, for lack of funds to travel abroad to international congresses, at informal seminars and meetings at home. Foreign funding agencies have made higher quality publications possible and have supported national surveys that are often cross-cultural and in English. Thus, there is a lack of surveys that provide a picture of local, regional and national surveys specificities (Nemchinov 1999).

Both the fundamental political changes and the difficult economic situation have led to a radical re-structuring of the social sciences. Within the first ten years (1989-1999) they have acquired a controversial status. Political science is in the privileged situation "of being the most varied and fast expanding field of research…Nevertheless, there is not much theory development, the bulk of the effort being spent on current events" (Nemchinov 1999, 102). This is part of a broader trend due to the 1989 transition to democracy and market economy: commercialization of education, the de-institutionalization of academic institutes and universities. Much of the social sciences is now driven by material needs. "Popular areas since the transition include polling, market analysis and survey research, often for foreign firms, academics and international institutions such as the International Monetary Fund (IMF)" and others (UNESCO 1999, 92). The UNESCO report comes to the dire conclusion that the shift from a single state source to multiple foreign funders has replaced cooperation with individual competitiveness, with the result that the "governance (and often the quality) of academic institutions has deteriorated significantly.." (UNESCO 1999, 92). If one compares the findings of the IZ report with those of UNESCO and OECD five years earlier, the situation does not seem to have changed substantially. The UNESCO report also takes a look at the state of the social sciences in Asia,

notably China and Korea. These countries have a long tradition of high regard for education, thus, the central governments are responsible for education. The higher education system is characterized by the separation of research and teaching. At various times and locations the requirements of public and foreign affairs have determined the establishment of subjects of scholarly attention as well as the development of universities and research institutions. Domestic goals have determined education programs. More recently there is growing interest in broadening education about the rest of Asia and the world. A negative aspect of the tradition in the education system is that it usually rewards seniority and mediocrity. There is also a strong emphasis on applied research. Nowadays this drives young people out of academia and into corporate research environments. The best social scientists now work in research institutes specialized to work on issues such as ageing, development or the environment.

As to international contacts the report states that "the systems suffer from introspection and isolation…In some places there is self-satisfaction, with elites talking only among themselves. Good work may be done, but it is not made generally available. In other places, the isolation may be physical rather than mental…In general, only a few elite academics, with strong English skills and, in the case of China, impeccable political credentials, are able to become members of the Asian academic jet set" (UNESCO 1999, 142p.).

The research issues in the social sciences in the region are the same as elsewhere – globalization, internationalization, protection of the environment. Research is theoretically oriented and strongly influenced by theories from the West. The same is true for the methodological orientation. Especially in China, research is focused on domestic issues and is carried out with a strong quantitative orientation. Although the social sciences have never flourished in Asia, they now fare better than ever before (UNESCO 1999, 143p.).

We have drawn on the reports on the social sciences in East Europe and Asia not only because the project is concerned with the SS&H in these regions but also because these reports point to the importance of the political environments including funding and the infrastructure for research. In other words, if the issue is to mobilize and to strengthen the SS&H in a particular region, it may be much more important as a first step – and ultimately more effective – to improve funding and the infrastructure (data bases, information services etc.) than to induce competition and call for international cooperation. High quality research cannot flourish without nourishment.

6) Recommendations for the assessment of the social sciences and humanities

The current rhetoric about excellence in science seems to streamline the unruly and inaccessible processes of knowledge production for policy decisions. In actual fact, however, the rhetoric only obscures the complexities of these processes, and the interventions based on the semantics may do more harm than good.[6] The following recommendations are based on the assumption that the objective is to assess the state of SS&H in the CIS and China in order to design measures that will help them become integrated more effectively into the global communities of social scientists and humanity scholars. This objective is quite different from very pragmatic evaluations and rankings of institutions that are now common in the EU and elsewhere. There is an evaluative side insofar as promising scholars and institutions have to be identified as targets for future measures of support and promotion. And there is a policy side insofar as these measures have to be designed to fit the specific conditions of the SS&H in their home countries. With respect to the latter we cannot claim specialized knowledge but will have to limit ourselves to fairly abstract principles.

1) The SS&H in their political environments – science policy for SS&H

The SS&H cannot thrive in environments in which open or even implicit censorship prevails and/or in which achievement in academia is heavily dependent on 'connections' to the predominant political class. Although it will not be possible to influence these conditions directly, the principle of separation of politics and science should be propagated loud and clear. This principle should find expression in funding patterns. State funding as well as funding by private donors should allow for independent basic research. At least 10% of the overall budget for SS and 20% of the budget for H should be set aside for non-programmed research.

A multiplicity of funding sources can also contribute to a plurality of theoretical approaches, thematic choices and methodological strategies. The same is true for a multiplicity of institutions where SS&H are represented. In countries where the size of the communities is so small that one cannot speak of a 'critical mass', it is crucial that scholars have contacts to colleagues outside their home country and that funds be available for exchanges of students and teachers alike. Thus, a first approach toward assessing the state of SS&H is to look at the size of the communities, the number of institutions where they are represented, and the funding patterns and the landscape of the funding institutions.

[6] We do not elaborate on the recent literature on the inadvertent, unintended consequences of evaluations on universities, scientists and their knowledge production. Cf. Whitley, Barker, Glaeser (eds.) 2007.

2) The range of quantitative indicators

Although the discussion about the applicability of quantitative bibliometric indicators is in flux and comes to different results with respect to social sciences and humanities, there is broad agreement that bibliometric indicators are insufficient by themselves to provide reliable assessments. However, used in conjunction with other indicators and descriptions they can serve an important purpose. For example, the sheer productivity in terms of number of articles or monographs can be seen as a predictor of quality in terms of citations. Consequently, any assessment of research institutions, university departments or smaller sub-units should try to obtain data on productivity and compare them with other institutions of similar size both nationally and internationally.

In the discussion of the differences between the SS&H and the natural sciences we identified national or language (i.e. cultural) orientation as one of the characteristics of the former. Nonetheless, especially the social sciences exhibit an increasing international networking with overarching research themes, international journals and scholarly associations. Visibility in international peer reviewed journals whose quality standards are established is one indicator of good international standing. However, the results must be controlled by the size of the national SS&H communities. For example, it may be that only a small number of individuals appear in these journals, representing a very small fraction of the particular national community. Such lack of visibility may have different reasons such as politically motivated limitations of access or resentment of international cooperation. Simple bibliometric indicators will not suffice to disentangle these diverse causes. Thus, publications in international journals – and perhaps cooperative authorships with international scholars – should not be taken as definitive indicators of quality of research but rather as relative, and above all as descriptors of just that: international visibility and integration. This means that they do not reflect the potential quality of work done in the national context and hidden from international view.[7]

External funds which are increasingly popular as quantitative indicators of research quality in some EU countries make hardly any sense in countries where these funds are very scarce and the procedures of their distribution are not transparent. Thus, before applying this indicator at all, the overall budget for SS&H and its sources must be known and it must be ascertained that the funds are distributed by a peer reviewed process open in principle to all members of the respective community of scholars.

[7] The bibliometric count of internationally co-authored publications (cf. Tables 13-24) shows the very small numbers for most of the CIS. There are also some interesting differences, for example in the extent of cooperation with neighbouring countries as compared to cooperation with Western countries. However, the numbers are by no means representative or reliable. They are the fraction of publications in the SS&H appearing in the SSCI and A&HCI.

Another indicator that gives a fairly reliable picture of the dynamism of university-based communities is the number of PhD (doctoral) students. Again the numbers have to be normalized by relating them both to the size of the national communities and to the number of supervisors with qualified academic training. (In some countries, especially in the developing South a large amount of teaching at universities is carried out by non-PhDs). Well functioning graduate and PhD programs indicate a sustained disciplinary culture.

3) Qualitative assessment of health and quality of SS&H research

Given the severe limitations of quantitative (mostly bibliometric) indicators in assessing the health and quality of SS&H research, it is unavoidable to take a direct approach. In a sense this will return the evaluative process to the scholars themselves without having to rely on proxies. We suggest, in this case, to resort to *two sets of criteria: organizational and intellectual.*

(1) Organizational criteria are both conditions of research as well as expressions of research culture. Thus, a healthy SS&H culture should have a sufficient size to allow for a plurality of approaches and methods (cf. below). Where countries are very small, close links to neighboring communities are desirable. Furthermore, the type of organization and overall role of SS&H in the universities as well as the existence of specialized centers must be considered. The crucial question is whether the SS&H have normal department status or if they are marginalized for example as adjunct courses in 'studium generale' programs. Connected with this is also the issue , where their students later find employment (e.g. in academia, as teachers, in industry, public administration or in the media?).

Furthermore, are the SS&H represented in national scholarly associations and professional societies? Such organizations serve, on the one hand, to strengthen a sense of identity, disciplinary unity and coherence as well as to formulate and uphold quality standards and to establish international contacts. On the other hand, they are also instrumental in representing the interests of the respective communities to the political authorities, the media and the public at large.

(2) Intellectual criteria are at the core of any assessment of the health and quality of a discipline or research field. SS&H do not have to be integrated into an international scholarly discourse to the same degree as the natural sciences in order to be qualitatively on a high level. Those research activities which are more narrowly focused on nation and culturally-specific subject matters and topics must be judged on their own merits. They must, above all, exhibit originality in their theories and methodologies. Indications for this are lively intellectual debates among the relevant scholarly communities, a recognizable progress of research over time, and, in the ideal case, an impact on public debates.

An important pre-requisite is the existence of independent quality-controlled (peer reviewed) scholarly journals and, especially in the case of the humanities, of more popular journals and/or print media catering to the intellectual elite of the country. SS&H that are entirely dependent on a few external sponsors or are only small inbred circles can hardly prove their value to civil society nor will they be open to intellectual stimuli from outside.

References

Cappellin, R., International knowledge and innovation networks for European integration, cohesion, and enlargement, *International Social Science Journal*, 180, 56, 2004, 207-225.

Carden, F., Tracking impact. Case studies on the social science – policy nexus, *International Social Science Journal*, 56, 179, March 2004, 135 – 152.

Cole, S., *Making Science. Between Nature and Society,* Cambridge:Harvard University Press, 1992.

Cole, J., Cole S., *Social Stratification in Science*, Chicago: University Press, 1973.

Engels, A., Ruschenburg, T., Weingart, P., Recent internationalization of global environmental change research in Germany and the U.S., *Scientometrics*, 62. 1, 2005, 67-85.

Frey, B. S., Evaluitis – Eine Neue Krankheit, Working Paper No. 293, June 2006.

Forbes, I., Abrams, D., International social science research: craft industry or baby behemoth?, *International Social Science Journal*, Vol. 180, June 2004, 227-244.

Heinze, T. et al., Identifying creative research accomplishments: methodology and results for Nanotechnology and Human Genetics, *Scientometrics,* 70, 1, 2007, 125-152.

Hicks, D. The difficulty of achieving full coverage of international social science literature and the bibliometric consequences, *Scientometrics*, 44, 2, 1999, 193-215.

Hollingsworth, R., Hollingsworth E.J., Major discoveries and biomedical research organizations: Perspectives on interdisciplinarity, nurturing leadership, and integrated structure and cultures, in: Weingart, P., Stehr, N., *PractisingInterdisciplinarity*, Toronto: University of Toronto Press, 1999, 215-244.

InformationsZentrum Sozialwissenschaften, Workshop 'The East European Social Sciences: Research Conditions and the Role of Information/Communication', IZ-Arbeitsbericht 33, Bonn 2004.

Iverson, A., Attribution and aid evaluation in international development: A literature review. May 2003. pdf version available under: http://www.idrc.ca/ev_en.php?ID=32055_201&ID2=DO_TOPIC (visited Dec. 20, 2006).

Larivière, V., Lebel, J., Lemelin, P., Collaborative Research in the Social Sciences and Humanities: Bibliometric Analysis of Practices, Observatorie des sciences et des technologies (OST), 2004, <http://www.ost.qc.ca/OSTE/pdf/rapports/2004/SSHRC.pdf> (visited Dec. 10, 2006).

Larivière, V., Gingras, Y., Archambault, E., Canadian collaboration networks: A comparative analysis of the natural sciences, social sciences, and humanities, *Scientometrics,*68, 3, 2006, 519-533.

Luhmann, N., Die Selbststeuerung der Wissenschaft, in: N. Luhmann, *Soziologische Aufklärung,* Vol.5. Opladen: Westdeutscher Verlag, 1970, 232-252.
Merton, R.K., The Matthew Effect in Science, in: R.K. Merton, *The Sociology of Science,* Chicago: University of Chicago Press, 1973, 438-459.
Nemchinov, V., The social sciences in the Russian Federation since the collapse of the Soviet Union, in: UNESCO, *World Social Science Report 1999,* Paris: UNESCO Publishing/Elsevier, 1999, 99-102.
Nicolaisen, J., Citation analysis. A*nnual Review of Information Science and Technology*, 41, 2007, 609-641.
OECD, *National Innovation Systems,* Paris: 1997;
http://www.oecd.org/dataoecd/35/56/2101733.pdf (visited Dec.2, 2006).
OECD, *Dynamising National Innovation Systems,* Paris, 2002;
http://www1.oecd.org/publications/e-book/9202071e.pdf (visited Dec. 2, 2006).
OECD, The Social Sciences at a Turning Point? Paris: OECD, 1999.

Schmid, K.P., Falsch gerechnet. Konjunkturforscher legen gern exakte Prognosen vor – und liegen meist daneben. Jetzt üben sie Selbstkritik., *Die Zeit,* 43, October 20, 2005.
http://www.zeit.de/2005/43/Prognose-Qual_?page=1 (visited Dec. 20, 2006).

Schoepflin, U.,Problems of representativity in the Social Sciences Citation Index, in:

UNESCO, *World Social Science Report 1999,* Paris: UNESCO Publishing/Elsevier, 1999.

Sternberg, R.J.(ed.), *Handbook of Creativity*, Cambridge: Cambridge University Press, 1999.
Subcommittee on Science, Research, and Technology, *National Science Foundation Peer Review*, 1. A Report of the Subcommittee on Science, Research, and Technology of the Committee on Science and Technology, U.S. House of Representatives, 94th Congress, Second Session, Washington D.C.: Government Printing Office, January 1976.
Stokes, D.E., *Pasteur's Quadrant. Basic Science and Technological Innovation*, Washington D.C.: Brookings Institution Press 1997.
Van Leeuwen, T., The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible, *Scientometrics*, 66, 1, 2006, 133-154.
Weingart, P. *Die Stunde der Wahrheit?Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft*, Weilerswist: Velbrück Wissenschaft, 2001.
Weingart, P: Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics,* 62(1), 2005, 117-131.

Weingart, P., Prinz, W.,Kastner, M., Maasen, S., Walter W., *Die sog. Geisteswissenschaften: Außenansichten*, Frankfurt: Suhrkamp.

Weingart, P., Sehringer, R., Winterhager, M., *Representations of Science and Technology*, Proceedings of the International Conference on Science and Technology Indicators, Bielefeld, Germany, 10-12 June 1990, Leiden: DSWO Press, 177-188.

Whitley, R. *The Intellectual and Social Organization of the Sciences*, 2nd edition, Oxford: Oxford University Press, 2000.

Whitley, R., Barker, K., Glaeser, J., (eds.), *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*, Yearbook Sociology of the Sciences, Dordrecht: Springer, 2007.

Wissenschaftsrat, *Empfehlungen zur Entwicklung und Förderung der Geisteswissenschaften in Deutschland,* Köln, 2005.

Annex:

Tables 1 – 12

Tables 13 - 24

Annex

Tab. 1-12:
Internationally visible publications of CIS

Armenia

Year	Publications		
	WoS	SSCI + A&HCI	SSCI + A&HCI%
1996	356	13	3,7
1997	318	3	0,9
1998	318	6	1,9
1999	338	6	1,8
2000	367	32	8,7
2001	372	5	1,3
2002	426	21	4,9
2003	452	3	0,7
2004	455	5	1,1
2005	449	7	1,6

Belarus

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	1207	21	1,7
1997	1312	10	0,8
1998	1255	35	2,8
1999	1287	14	1,1
2000	1165	20	1,7
2001	1016	11	1,1
2002	1081	18	1,7
2003	1043	16	1,5
2004	1026	21	2,0
2005	1035	13	1,3

Kazakhstan

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	222	11	5,0
1997	190	7	3,7
1998	215	14	6,5
1999	198	5	2,5
2000	222	22	9,9
2001	201	9	4,5
2002	222	7	3,2
2003	264	9	3,4
2004	254	9	3,5
2005	228	12	5,3

Azerbaijan

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	219	11	5,0
1997	187	1	0,5
1998	183	2	1,1
1999	143	1	0,7
2000	165	1	0,6
2001	149	0	0,0
2002	195	3	1,5
2003	251	4	1,6
2004	227	6	2,6
2005	251	1	0,4

Georgia

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	241	26	10,8
1997	254	5	2,0
1998	233	3	1,3
1999	295	10	3,4
2000	303	31	10,2
2001	252	10	4,0
2002	326	24	7,4
2003	265	12	4,5
2004	339	11	3,2
2005	366	14	3,8

Kyrgyzstan

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	28	3	10,7
1997	24	2	8,3
1998	37	4	10,8
1999	36	2	5,6
2000	44	6	13,6
2001	58	4	6,9
2002	51	7	13,7
2003	26	2	7,7
2004	41	6	14,6
2005	64	6	9,4

Moldova

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	199	5	2,5
1997	196	0	0,0
1998	171	2	1,2
1999	185	5	2,7
2000	189	8	4,2
2001	177	0	0,0
2002	183	1	0,5
2003	215	3	1,4
2004	173	3	1,7
2005	232	2	0,9

Tajikistan

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	40	2	5,0
1997	42	0	0,0
1998	26	0	0,0
1999	44	0	0,0
2000	32	3	9,4
2001	36	1	2,8
2002	45	0	0,0
2003	32	1	3,1
2004	32	0	0,0
2005	32	0	0,0

Ukraine

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	4240	102	2,4
1997	4269	36	0,8
1998	4378	37	0,8
1999	4591	125	2,7
2000	4459	107	2,4
2001	4371	37	0,9
2002	4301	56	1,3
2003	4059	38	0,9
2004	4177	61	1,5
2005	4310	56	1,3

Russia

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	28756	1017	3,5
1997	29670	1042	3,5
1998	29128	1141	3,9
1999	28657	1008	3,5
2000	28962	1149	3,9
2001	26800	720	2,7
2002	27726	905	3,3
2003	26718	811	3,0
2004	26882	825	3,1
2005	26661	723	2,7

Turkmenistan

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	16	2	12,5
1997	12	1	8,3
1998	9	0	0,0
1999	5	0	0,0
2000	14	2	14,3
2001	10	2	20,0
2002	10	0	0,0
2003	6	0	0,0
2004	5	1	20,0
2005	5	0	0,0

Uzbekistan

Year	Publications		
	WoS	SSCI+ A&HCI	SSCI+ A&HCI%
1996	353	6	1,7
1997	382	1	0,3
1998	343	2	0,6
1999	343	3	0,9
2000	357	5	1,4
2001	342	5	1,5
2002	372	4	1,1
2003	361	7	1,9
2004	363	7	1,9
2005	366	2	0,5

Tab. 13 - 24: Share of internationally co-authored publications of CIS

Armenia

Publications (Total: 96)		**Coauthorship** (> 1%)
Abs.	**Perc.**	
16	17%	All countries
9	9%	USA
3	3%	Italy

Belarus

Publications (Total: 179)		**Coauthorship** (> 1%)
Abs.	**Perc**	
50	28	All countries
16	9	Russia
11	6	Germany
10	6	USA
7	4	UK
5	3	Ukraine
5	3	France
4	2	Hong Kong
4	2	Netherlands
4	2	Canada
3	2	Taiwan

+ 10 other countries

Kazakhstan

Publications (Total: 104)		**Coauthorship** (> 1%)
Abs.	**Perc**	
26	25	All countries
9	9	USA
6	6	Russia
6	6	UK
3	3	Germany
2	2	Uzbekistan
2	2	Italy
2	2	Kyrgyztan

+ 13 other countries (1 P)

Azerbaijan

Publications (Total: 29)		**Coauthorship** (> 1%)
Abs.	**Perc**	
6	21%	All countries
2	7%	Russia
2	7%	UK
2	7%	Ukraine
2	7%	USA
1	3%	Kazakhstan

Georgia

Publications (Total: 146)		**Coauthorship** (> 1%)
Abs.	**Perc**	
42	29	All countries
22	15	USA

11	8	Germany
11	8	UK
6	4	France
6	4	Switzerland
6	4	Hungary
6	4	Russia
4	3	Estonia
4	3	Israel
4	3	Netherlands
4	3	Poland

+ 24 other countries

Kyrgyzstan

Publications (Total: 42)		**Coauthorship** (> 1%)
Abs.	**Perc**	
13	31	All countries
5	12	USA
2	5	Uzbekistan
2	5	Kazakhstan
2	5	Turkey
1	2	France
1	2	Israel
1	2	Japan
1	2	China
1	2	Russia
1	2	Tajikstan
1	2	UK

Moldova

Publications (Total: 29)		**Coauthorship** (> 1%)
Abs.	**Perc**	
7	24	All countries
3	10	Romania
2	7	UK
1	3	France
1	3	Latvia
1	3	Lithunia
1	3	Macedonia
1	3	USA

Tajikstan

Publications (Total: 7)		**Coauthorship** (> 1%)
Abs.	**Perc**	
4	57	All countries
3	43	Russia
3	43	USA
1	14	France
1	14	Israel
1	14	Kyrgyzstan
1	14	Poland
1	14	Ukraine
1	14	Uzbekistan

Turkmenistan

Publications (Total: 8)	**Coauthorship** (> 1%)

Abs.	Perc	
3	38	All countries
2	25	Turkey
1	13	Denmark
1	13	Israel
1	13	Kazakhstan

Russia

Publications (Total: 9341)		**Coauthorship** (> 1%)
Abs.	**Perc**	
1277	14	All countries
626	7	USA
223	2	UK
152	2	Germany
94	1	France

+ 82 other countries

Ukraine

Publications (Total: 655)		**Coauthorship** (> 1%)
bs.	**Perc**	
151	23	All countries
59	9	USA
36	5	Russia
23	4	UK
25	4	Germany
12	2	Netherlands
10	2	Austria
9	1	Sweden
7	1	Canada
7	1	Poland
7	1	France

+34 other countries

Uzbekistan

Publications (Total: 42)		**Coauthorship** (> 1%)
Abs.	**Perc**	
16	38	All countries
6	14	USA
3	7	France
2	5	Canada
2	5	Kazakhstan
2	5	Kyrgyztan
2	5	UK
2	5	Ukraine
1	2	Bangladesh
1	2	Cambodia
1	2	Japan
1	2	Netherlands
1	2	China
1	2	Rusia
1	2	Tajikstan