ERC Working Paper in Economic 03/11
December 2003

Difficulties and Trade-offs in Performance Evaluation in Social Sciences A Turkish Perspective

Fikret Şenses
Department of Economics
Middle East Technical University

Economic Research Center
Middle East Technical University
Ankara 06531 Turkey
www.erc.metu.edu.tr

DIFFICULTIES AND TRADE-OFFS IN PERFORMANCE EVALUATION IN SOCIAL SCIENCES - A TURKISH PERSPECTIVE

FİKRET ŞENSES[1]

DEPARTMENT OF ECONOMICS

MIDDLE EAST TECHNICAL UNIVERSITY

1) INTRODUCTION

Increasing the performance of academic institutions and their staff has in recent years emerged as an important objective of university administrations not only in Europe and the United States but also in some developing countries such as Turkey. The emphasis on performance has been guided primarily by increased importance attached to the efficiency of resources allocated to higher education by central governments. It can also be linked to the emergence of the market-oriented university which is required to adhere to strict performance criteria for access to its already reduced share of public funds. In this interpretation, the emphasis on performance evaluation is considered as a part and parcel of the neoliberal approach to institutions of higher education and serves the purpose of managing and controlling their activities[2].

Research output of academic staff has headed the list of criteria used for evaluating performance. To increase the research output of their academic staff, universities have begun to offer a set of incentives, including the reduction of course load and provision of a variety of pecuniary incentives, such as support to facilitate conference participation.[3] Likewise, in some countries the allocation of research funds among different universities and within various departments of the same university is based on research performance of academic staff as measured by the number of publications weighted according to where the publication is made.[4] A similar criterion is used in the ranking of universities in the United States by various organizations. This ranking is taken as an indication of the quality of these institutions, most notably by applicants for graduate study. The fact that there exists a transfer market for academics with quality publications, not too different from the one for successful sportsmen, is an indication of how seriously universities take their ranking. In Europe, too, the past two decades have witnessed most countries engaging themselves with a similar endeavor with the United Kingdom perhaps going further than others in this direction.

In Turkey, the past decade or so has seen an increasing quality consciousness by universities most notably by those in which tuition is in English. The increased emphasis on quality has been accompanied by the adoption of various assessment schemes by these universities which have invariably attached top priority to the research and publication

---

[1] This paper was inspired by the workshop on the Evaluation of the Performance in Social Sciences held at the Middle East Technical University in Ankara on 26-27 June 2003 in whiich I participated as a session chair and commentator. I thank Prof. İlhan Tekeli with whom I had a chance to discuss the issues involved on this and other occasions for encouraging me to write this paper.

[2] See Fish (2003) for an elaboration of this view.

[3] See Laband and Tollison (2003: 161-62).

[4] For the use of academic performance criteria in the allocation of research funds in Australia, for example, see Butler (2003:150).

performance of their academic staff. Two closely interrelated factors have played a major role in this process. First, in order to increase their national and international standing universities have sought accreditation by an international body. The second factor has been the imposition of stricter criteria for academic appointments and promotions.

Although there has been a broad agreement on the objective of a more efficient utilization of resources, the selection of criteria for performance evaluation has emerged as a highly controversial subject. The controversy has so far been dominated by several key issues. Should research output preserve its dominant status on the list of evaluation criteria? If so, what should be the instruments of evaluating research output? What are the main difficulties and trade-offs involved when developing countries adopt the performance criteria of developed countries? Is research output published in a journal with international acclaim inherently and necessarily more valuable than its domestic counterparts?[5] Should performance evaluation criteria be somewhat different in developing countries such as Turkey?

The main objective of this study is to briefly examine the process of performance evaluation in social sciences in Turkey from an international perspective and pinpoint the main difficulties and trade-offs involved in this process and suggest new ways of thinking to overcome them. Our discussion will be confined to social science research and publications by universities which in any event account for the bulk of these activities. Evaluation in social sciences can generally be conducted at the institutional, project and personal levels. While also touching upon the others, our emphasis will be on evaluation at the level of the individual researcher. We shall lean heavily on cases drawn from the field of economics and the Middle East Technical University with which we are more familiar.

The plan of paper is as follows. After this introduction, in Section 2 of the study we examine the pattern of evaluation in Turkey against the background of cases in other countries. This is followed in Section 3 by a critical evaluation of the performance criteria used in Turkey and draw attention to the difficulties and trade-offs involved. Finally, Section 4 is devoted to our recommendations, drawn as possible solutions to the problems identified in the preceding section.

2) RECENT TRENDS IN EVALUATION PRACTICES- AN OVERVIEW

The discussion on the evaluation of performance should consider the question of whether we should measure performance at all as its starting premise. In the absence of concern of central governments with the efficiency of public resources allocated to higher education and selecting appropriate criteria for academic appointments and promotions much of the arguments for assessing performance would lose their power. For there is more or less an automatic mechanism of assessment in the academic community. Without the painstaking efforts of quantifying performance indicators a simple survey of members of the academic community could easily identify the successful institutions and individual researchers as well as the laggards. The performance of the great social scientists of the past two centuries, for instance, has perhaps never been assessed, but their work is still alive and well.

The increased emphasis on the quality of higher education by governments as well as the users of higher education services has in a number of countries led to a proliferation of various organizations monitoring quality and accrediting the institutions of higher education that pass their test. Although this is done by government organizations in most countries, as

[5] See Tekeli (2003 b: 1) on this point.

many as 19 nongovernmental accrediting organizations are active in the United States monitoring and reviewing the quality of colleges and universities.[6] In the same vein, the emphasis on quality in the European context as encouraged by the Bologna Declaration in 1999 has led to the establishment of the European Network for Quality Assurance in Higher Education (enqa)[7] with a view of increasing co-operation in performance evaluation among European countries.

The growing emphasis on evaluation of performance since the early 1980s has not led to the emergence of a "best practice". Instead, assessment criteria in various countries and the mechanism of implementing them have shown a great deal of variation across countries and over time. In the United Kingdom, for example, in the evaluation of different university departments, main emphasis has been on "publications, research culture and research organization" while in the Netherlands these criteria have been based on "quality, productivity, and relevance and viability". Denmark has, on the other hand, adopted a more qualitative assessment procedure based on information gathered through visits to individual departments.[8] In the Netherlands, the organization and implementation of assessments have been conducted by the Association of Universities while in the United Kingdom and Denmark the Ministry of Education and Ministry of Research have, respectively been in the driver's seat.

In recent years, evaluation processes in Turkey, particularly those pertaining to academic promotions, have increasingly emphasized publications. The Higher Education Council compiles information from all universities on the publication performance of their academic staff, clearly distinguishing between domestic and international publications and publishes the results each year. It implicitly encourages international publications and takes the rise in their number as a strong sign of academic progress.

Rules and regulations guiding the promotion to associate professorship, for instance, which relied on candidates to write a thesis or long paper as well as testing them on their ability to lecture was changed in 1982, with the new rules relying primarily on publication performance to be followed by an oral examination in front of the same jury. The new set of rules introduced in 2000 has reinforced this set of preferences, requiring candidates to accumulate a minimum number of total points from their publications and assigning highest points to international publications.

The Middle East Technical University has strict criteria for academic appointments and promotions which require the candidate to have as a prerequisite a minimum of two international publications for promotion to the rank of associate professor and another two for promotion to full professorship. Some other universities have similar criteria with some actually resorting to international referees for certain appointments which by discarding altogether publications in the Turkish language take the emphasis on international publications to new heights.

The Turkish Academy of Sciences has a scheme which offers pecuniary rewards for international publications such as papers published in SSCI journals and books and chapters in books published by international publishers. There is a long list of journals divided into several categories according to their impact levels with papers published in them rewarded on different scales. More significantly, the annual awards by the Academy of Sciences for

---

[6] These organizations are reviewed by the United States Department of Education or the Council for Higher Education Accredition (CHEA) whose database contains information on 6,500 US institutions of higher education. See http://www.chea.org.
[7] See http://www.enqa.net
[8] See Siune (1998: 5).

contributions to academic community, and academy membership itself are conditional on performance evaluated on the basis of the metric of international publications and citations. The Middle East Technical University along with several other universities also offers pecuniary rewards for such publications. Its criteria for the utilization of subsidized university accommodation for an extended period include academic performance which again reflects heavy emphasis on international publications. In some other universities the reward for international publications takes the form of giving grants for participation in international conferences. In none of these cases mentioned above is there a reward for publications in the Turkish language whatever their quality and overall impact. This emphasis on international publications which has been justified on grounds that it accords with "universal quality criteria" has been challenged by some observers.[9]

3) DIFFICULTIES AND TRADE-OFFS IN EVALUATION

The main difficulties and trade-offs involved in the current evaluation processes can be grouped under several closely interrelated headings.

Neglect of Education

The first pertinent question that arises in the context of evaluation in social sciences is whether assessments should be based solely or even primarily on research and publication performance. In sharp contrast to the current practice, there is a need, especially in developing countries like Turkey, to consider this performance together with performance in education as integral and mutually reinforcing parts of the same process.

Much of economics that is taught in developing countries, for example, is almost a perfect replica of its counterparts in developed countries, as shaped and dominated by the preferences of the United States. There has been a clear and strong tendency of curricula in the United Kingdom and continental Europe taking their cue from the American universities, especially in the design of their graduate programs. A similar tendency has been observed in Turkey as encouraged by the increase in the number of universities in which tuition is in a foreign language, predominantly English. [10] The curricula in these universities in particular have borne such a close similarity with their American counterparts that in some cases even course code numbers and text books are identical. It should therefore be recognized at the outset that social science research does not take place in a vacuum but is instead strongly influenced by the dominant paradigm of the day and the content of education as reflecting the biases of developed countries.

The dominant paradigm in economics has been neoclassical economics, notorious for its neglect of the historical, cultural, and political context within which economic phenomena takes place. Its relevance to real life problems is questionable even in the developed economies let alone the developing ones. As a reflection of the biases of economics education, scientific journals in the field are full of articles purporting to deal with different facets of economic life without even a touch of historical and political analysis and any real effort to cooperate with other social sciences in the form of interdisciplinary studies. Under these circumstances, attempts to introduce political arguments into economic analysis often face the charge from mainstream economists that such arguments are unscientific.

---

[9] See, for example, Şenses (1993:180) and Tekeli (2003).

[10] For an overview of the main problems of economics education in Turkey see Şenses (1986).

The main claims on the time of the university professor come from education and research with a clear and strong trade-off between them. The research-active professor might be well placed to transmit his research effort into teaching and thereby contribute to the quality of education. Given the strong emphasis on research in most evaluation processes, as appropriately summarized in the well-known phrase "publish or perish" however, it is usually the quality of education that suffers. Additional time devoted to research would necessarily mean less time allocated to students' requirements. A research-active professor, especially if he is also an active member of the international conference circuit might increase the ranking of his university by several positions, but this might mean cancelled office hours and rescheduled classes if not missing them altogether. More importantly, this might denigrate education as an unrewarding activity which one does as a matter of routine rather than with a deep sense of duty. Reduced accessibility to students would mean not only less time spent on commenting on their essays, listening to and advising them on their career plans, fewer reference letters written, reduced familiarity with their problems and their talents, and in the final analysis a missed opportunity to act as a role model for young people at a crossroads in their careers. Although the trade-offs involved are clearly substantial, one cannot easily quantify them.

Not all the costs of neglecting education in this way are incurred at the individual level either. From the point of view of the development of a domestic scientific community, it is important not to overlook the fact that recruitment to the academic profession sometimes takes place through the diligence of the dedicated teacher at a surprisingly early stage of the student's education. Teaching is important in giving students research orientation and in attracting their attention to important and socially relevant research spheres, not to mention in developing their research skills. An undergraduate essay properly supervised may lay the foundations of a research paper with wide acclaim with a lag of say ten years.

Problems of Quantitative Assessments

An evaluation process based entirely or even primarily on research and publication performance with emphasis on numbers is also surrounded by a number of problems, headed by the neglect of quality considerations. The quantification of performance on the basis of preannounced criteria has often been defended on the basis of objectivity. Although this defense can be justifiable in countries like Turkey with a tradition of cronyism this approach is not without loopholes. Even when a set of performance criteria is agreed upon, their implementation may pose difficulties arising basically from efforts of those assessed to get round the rules. When the maximand is the number of publications, papers and books may, for example, be subdivided into separate publications and thus tending to get shorter. In the United Kingdom, it has been observed that the transfer market for successful academic staff with a high publication rate gets livelier just before the beginning of the next round of evaluation as some departments want to boost their standing by such back door tactics.[11] Quantitative criteria are not altogether free from subjective interpretation, either. In the Turkish case for example, candidates for promotion to the rank of associate professorship may find, for example, that the points they have accumulated from their articles in refereed journals vary from one jury member to another as some of the latter may simply not accept the allegedly "refereed" status of a certain journal.[12]

[11] See Siune (1998:7).
[12] Another example of the difficulty of enforcing predetermined evaluation criteria comes from Belgium where there were cases of academics pretending that they were conducting joint research whereas their real intention was to get the funding designed to increase such research activity. See van Langenhove (1998: 36)

Research activity is, by its very nature not very conducive to being confined within the straitjacket of a predetermined set of criteria, especially when they are expressed in terms of sheer numbers. The trade-offs here involve the possibility that such guided research activity may, by pressurizing the researcher to maximize his publications before the next assessment round, remove much of the excitement and surprise element from this activity and may entail loss of originality. Likewise, there is the possibility that his *magnum opus* would be treated at par with other work of much lesser quality and impact. Among lesser problems in this respect is the arbitrariness in accounting for individual contributions in joint research output with the current practice varying from the paper apportioned equally amongst the authors to each author receiving full credit for one whole paper.[13]

Problems of Qualitative Assessments

The problems abound when one goes beyond the quantity of publications to assessing quality which is not an easy task not least because there has been a vast increase in the degree of specialization in all social sciences. Once the problem of finding the right assessors is overcome, there is the problem of assessor prejudices and biases bedeviling an objective assessment.[14] Even when the assessor can successfully free himself from the social dimension, it is known that one tends to treat academic work with which one is familiar more leniently than others. Even the age of the assessor may introduce a degree of subjectivity to assessments.[15] When it is a committee of assessors making the evaluation as in the case of assessments pertaining to promotions to the rank of associate professorship in Turkey there is a certain degree of arbitrariness introduced into the system with the quality standards often showing a great deal of variation from one committee to the other.[16]

Another hurdle in this respect pertains to the different perspectives of the assessed and the assessors rendering it difficult to reach a generally accepted set of criteria that survives over a period of say ten years. The assessors are generally inclined to move towards stricter evaluation criteria to such an extent that some of the assessors who have served in committees ruling on promotions to the rank of associate professorship and others in Turkey, for instance, may not fulfill the criteria which they find themselves in a position to enforce. When international assessors are employed these problems are aggravated by the language barrier and the bias against publications in the native language.[17] What complicates the problem further is that the real value of a publication can be understood only after a considerable time lag and/or after it is being used by another researcher.[18]

Problem of assessors is as important as the criteria for assessment. The first problem here stems from the biases and prejudices of assessors and pressures of cronyism on them that may get in the way of objective assessment. On the basis of my experience in this task so far, I can

[13] See. Butler (2003:145) on this point. At the Middle East Technical University, each author is credited with 80 percent of the paper.

[14] A study in which respondents were asked to rank various journals according to their significance has found that Journal of Economic and Statistical Theory and Journal of Regional Studies and Economic Change were ranked respectively in the top third and bottom third of all journals, although none of the two journals existed in reality. This no doubt was a reflection of respondents' subjectivity in attaching more significance to journals which seemingly had a theoretical and quantitative content. See. Brauninger and Haucap (2003: 178).

[15] While German economists over the age of 55 ranked Kyklos fourth among all economics journals, those below the age of 35 ranked it 38th. See Brauninger and Haucap (2003: 189).

[16] It seems that this is also a common problem faced in other countries. For example, a study by the National Science Foundation in Denmark has shown that 25% of research proposals moved from the category worthy of support to refusal or vice versa, when proposals were re-refereed. See Hansen (1998: 70).

[17] See Newton (1998: 17).

[18] See Brauninger and Haucap (2003:175).

only say that this is not an easy problem to overcome. There is also the additional problem of the shear volume of the material to be examined. In the case of promotions to the rank of associate professor in Turkey, for example, each of the five centrally appointed jury members are required to go through independently the dossier of up to five candidates who include all the relevant information about themselves together with a copy of all their publications. The jurors are expected to assess all this and complete their individual reports on each candidate in a period of two months which is no mean task.

There are other trade-offs involved in orienting the evaluation process towards quality. The time spent on reviewing others' work and the burden that this imposes on the time of a professor may be immense. In Turkey, for example, professors are required to review the publications of as many as five candidates seeking promotion to the rank of associate professor. The fact that the same professor is required to review books and referee a large number of papers for academic journals and that such demands are usually bigger for the research active professor may create important trade-offs in terms of lost research output.

Emphasis on International Publications

In view of the difficulties of objective assessment of quality especially at a time of rapid growth of publications, the trend has been to search for an indicator on which broad agreement can be reached. The indicator selected for this purpose has been books and chapters in books published by reputable publishers in developed countries, and most importantly articles published in journals included in the Social Science Citation Index. Such a selection can be justified on several counts. Internationally reputable publishers are renowned for strict evaluation of proposals and stringent refereeing and editing of books prior to publication. Likewise, the articles published in SSCI journals are usually selected after they pass a strict process of refereeing with stringent requirements. They are expected to have a significant research question and an adequate analytical content with an appropriate methodology and data analysis, and contribute to the state of knowledge in the field, properties which some of the non-SSCI and native journals may lack. Emphasis on international publications with high academic standards may, by increasing the aspirations of developing country researchers be instrumental in raising the quality of their work. International publications may also increase the interaction of domestic researchers with the international academic community which may lead to fruitful cooperation.

These positive contributions of an evaluation process based on international publications notwithstanding, there are a number of difficulties and trade-offs which need to be emphasized.

First, international publications have been increasingly associated with books and articles published in the English language predominantly in the United Kingdom and the United States. This is especially so with articles published in SSCI journals which probably have an even stronger Anglo-American bias. Journals published in developing countries, including Turkey[19] are on the whole not covered by this index. In view of the parochial structure of SSCI, one must guard against ascribing articles published within this framework "universal" status. Another fundamental problem in this respect is associated with the fact that the sharp increase in the number of scientific journals in social sciences has not been accompanied by a corresponding increase in their impact as measured by the number of citations articles

[19] Only two Turkish journals have been covered by this index during the 1973-99 period. These are the Turkish Journal of Pediatrics and Türk Psikoloji Dergisi (Turkish Journal of Psychology) (included in 1994). See Yurtsever et al (2001).

published in these journals receive. [20] In the field of economics, too the proliferation of academic publications has failed to create the expected impact. For example, the sharp increase in the number of scientific journals was accompanied by a large number of published articles not receiving any citations. Of all the economics articles published in SSCI journals in 1996, 26 percent received no citations during the five years after their publication, a massive 85 percent received ten or fewer citations.[21] One suspects that the pressure on academics to publish has led to an overgrowth of publications beyond what can be justified by social need and social value considerations.[22] It seems that putting primary emphasis on academic research output on efficiency grounds has paradoxically led to inefficiency through overproduction. A similar trend can also be observed in Turkey where the emphasis on publications in academic appointments and promotions has led to a proliferation of journals, published by various universities, foundations, and financial institutions with obscure review criteria for papers submitted for publication. The dire approach of taking the number of international publications as the maximand in evaluation processes is likely to have delayed the development of scientific journals of high quality at home.

Second, emphasis on international publications carries with it the potential risk of diverting domestic research effort away from local concerns and indigenous research agendas. Admittedly, defining what is local and relevant research has become increasingly more difficult. Does it, for example, refer to one's own community, province, region, country or the region or the continent in which one's country is located?[23] Drawing the line between these categories has become rather blurred especially during the past two decades of rapid globalization. It is also possible that an international publication may be more responsive to and reflect a deeper understanding of local issues than some of its domestic counterparts. These difficulties notwithstanding, some broad agreement can still be reached on what is local, especially from the perspective of developing countries. Not only do these countries face some common problems, each has special problems of its own which may provide the basis of socially relevant local academic research. Such a reorientation may be necessary as a safeguard against the risk of such pressing issues being eclipsed by an overemphasis on international publications. We may, therefore, expect social science research to be focused more on domestically relevant issues with its impact confined again mostly within the domestic sphere.[24] It is precisely at this point that giving precedence to international publications in evaluation processes may cause the most serious difficulties and trade-offs.

The risk of the emphasis on international publications diverting research effort away from local concerns is highest for research projects financed by foreign foundations, foreign research institutions and international bodies with multi-billion-dollar research budgets, creating unfair competition for their domestic counterparts with meager resources. The institution that provides the funding of the research project determines in the first place the research topic which more often than not reflects its own priorities and preferences. The influence of the financing institution may in some cases extend well beyond the research subject into the sphere of the formulation of specific research questions, and the design,

---

[20] For example, despite the fact that Australia's share in the total number of articles in SSCI journals in economics has increased by 25%, its ranking in terms of the impact of these publications among 11 OECD countries has declined from sixth to tenth. See Butler (2003: 143).

[21] See Laband and Tollison (2003: 162-64).

[22] See Laband and Tollison (2003: 165).

[23] I owe this point basically to a comment made by Prof. Ahmet Acar at the workshop by which this paper was inspired.

[24] See Tekeli (2003b:7) for an elaboration of this point.

methodology and content of the research project.[25] More seriously, the research project may even be designed to convey and even propagate the message of the donor to the domestic research and policy-making circles. In international projects involving a number of country case studies, individual studies bear a close resemblance with each other with the most significant novelty being the otherwise identical tables simply filled up with country-specific statistical information. Some positive spillovers from such research activity such as increased contact with researchers from other countries and generation of some new knowledge notwithstanding, one should not overlook the social costs involved.

International publications not involving foreign- financed projects are also associated with similar risks, albeit to a lesser extent. Although there are a large number of SSCI journals specializing in the problems of developing countries, some of which tackling problems primarily from the point of view of these countries, emphasis on publications in these journals may again divert domestic research agendas to the particular areas of interest of these journals. The amount of attention that developments in a particular country attracts in the international community such as the devastating earthquakes or the intermittent and deep economic crises that occurred in Turkey in recent years may increase the chances of publication in an international journal. In such cases, articles published in these journals may have a highly informative as opposed to an analytical content, with limited contributions to knowledge beyond what is already available domestically. The detailed treatment of certain aspects of the Turkish economic crisis which would be regarded as an indispensable part of a paper written in Turkish for a Turkish language journal would be regarded as unnecessary detail for the international audience.

A detailed examination of international publications by Turkish scholars is beyond the scope of this paper. However, one gets the impression that a good many of them consist history, economics, psychology, and sociology papers pertaining to Turkey[26]. In many cases they are also informative in nature. There is a remarkable lack of even comparative studies let alone those dealing with general issues of universal interest. Papers which are not simply informative fall in the sphere of applied research, testing, for example, a hypothesis tested elsewhere for the Turkish case. There is a striking lack of papers which can be regarded as truly path-breaking with original research questions and/or representing a methodological breakthrough. Under these circumstances it would be a gross error to rank an international publication more favorably than its domestic counterpart.

Another problem associated with relying on international publications as the main criterion for performance assessments arise from the restrictive stance of international journals not only in terms of research topics but also in the selection of appropriate methodology. For example, the penetration of quantitative techniques into the field of economics has taken place with such speed that it has been inundated with journals specializing in the use of such techniques, effectively barring papers with a political economy stance.

Emphasizing international publications may also increase the segmentation among social science practitioners, especially in developing countries. The emphasis on international

[25] A more recent trend in this respect has been research institutions and foundations based in foreign countries and international organizations conducting joint research with a native scholar/institution. In many cases the role of the native counterpart is reduced to supplying information, arranging interviews, and conducting the research related to the country in question without much say in developing the main research questions and the analytical framework. This no doubt limits the amount of positive spillovers from such activity to the domestic academic community.

[26] We have found that out of a total of 1023 papers published by Turkish scholars in SSCI journals during the 1990-99 period as given in Yurtsever et al (2001), a massive 33.7 percent had a word such as Turkey or Turkish in the title suggesting in some significant way connection with Turkey.

publications in academic promotions and appointments has no doubt created an unfair advantage for universities with foreign language tuition and academics trained in foreign countries. Such an emphasis hinders effective communication between those who take their main point of reference as international publications such as the SSCI journals and those taking it as the journals published in the native language. In practice, this, apart from erecting formidable barriers of communication among academics of different higher education institutions, precludes the development of scientific terminology in the native language, further hinders the development of an effective academic community, and also creates fertile ground for brain drain.

To overcome the difficulties of relying solely on the number of publications, assessments have increasingly relied on the impact of an article as measured by the number of citations it receives in important international journals, usually the SSCI journals. This approach is also bedeviled by a number of difficulties. The accumulation of knowledge is a highly incremental process and identifying the actual source of a scientific contribution is particularly difficult in social sciences. There is always the risk of a reference being misaddressed to someone who simply recites it rather than the original contributor.[27] Simply counting the number of citations also fails to take account of the nature of citation and makes no distinction between references acknowledging the contribution of a previous paper in developing the analytical content of the paper from another reference simply citing the previous studies on the subject. Even when one accepts the number of citations as a yardstick for evaluations, the big differences among international journals in terms of their quality and impact which may also show considerable variation over time[28] are likely to present additional problems. The exclusion of citations in books from this metric is a major drawback.

Other Difficulties

Another difficulty in assessment arises from the fact that a uniform set of evaluation criteria may fail to come to grips with the intricacies of individual disciplines in social sciences. The Middle East Technical University for example has divided all disciplines into only three categories, basic sciences and engineering, architecture, and social and administrative sciences,[29] amalgamating thereby in the latter category fields as divers as political science, sociology, and economics under one broad heading. Likewise, subjecting all institutions of higher education to a uniform set of criteria would also be problematic as they differ considerably in terms of their date of establishment and the level of resources they command.

While emphasis on performance assessments may have some positive effects such as increasing the diligence and productivity of academic staff it involves serious trade-offs in a number of other spheres. The proliferation of different evaluation schemes in different countries during the past two decades, increased complexity and bureaucratization of evaluation schemes not to mention the considerable financial burden it has imposed on central

[27] See Tekeli (2003b: 8) on this point. To support this view at an anecdotal level, my effort to trace the origins of an argument linking corporate profits to capital accumulation in the Korean context to early development thinking of the early 1950s has met with the remark by one of my doctoral students that the origins of the argument may go back to Schumpeter's idea of creative destruction.

[28] For example, Brookings Economic Papers which was ranked 137th in 1998 among 161 journals in economics jumped to 3rd rank in 2000 on the basis of the impact factor of journals, defined as the ratio of the number of citations that articles published in a journal during a certain period to the number of articles. See Brauninger and Haucap (2003:176).

[29] See http://www.po.metu.edu.tr

governments.[30] The incentive effects of performance assessments should be balanced against the demoralizing effects on those whose assessment results do not measure up to the standards of the department. This may lead to a sense of letting the team down or worse still to an environment of “rat race” where unproductive competition takes the place of fruitful cooperation which forms the basis of an efficient academic community. Moreover, expecting the same performance from all individuals and academic units may discourage the laggards which may be costly especially in developing countries where emphasis should also be on recruiting people to academic profession.

4) WHAT IS TO BE DONE?- CONCLUDING REMARKS

The evaluation of performance in social sciences is not independent of the broader issue of the evolution of universities. What society demands from universities and how universities respond to those demands over time have inevitably affected the meaning attached to quality and even efficiency issues. Throughout the long period of evolution of universities during the past several centuries emphasis on education and research has preserved its role at central stage. In the long history of higher education, the emphasis on evaluation of performance using an increasingly complex array of criteria is a relatively recent issue going back only to the past two decades.

The current schemes of evaluation of performance in social sciences is bedeviled by a number of serious difficulties and trade-offs especially in developing countries which are trying to meet the formidable challenge of creating an academic basis rather than facing these problems from a position of strength as in developed countries with academic traditions going back several centuries. The common reaction of analysts facing qualitative factors has been to move towards indicators that are measurable. The overemphasis on the quantity of publications at the expense of quality has, however, resulted in a lot of the research output being unused. Such emphasis on numbers has not only failed to measure what we intend to measure but has also created important trade-offs, especially in the setting of developing countries. This presents a strong challenge to existing evaluation schemes and underlines the need for searching for alternative mechanisms.

From a developing country perspective, the main pillars of the new approach to the evaluation of performance in Turkey should consist of the following interrelated considerations.

First, education performance should along with research be a major component of the evaluation process which should recognize the close and mutually reinforcing interaction between these two activities at all levels of higher education. Text-book writing is at present perhaps the least rewarding of all evaluation criteria. The text books currently used in developing countries even when they are not in a foreign language do not in many cases go far beyond translating books published by international publishers into the local language without even bothering to introduce elements of the local context. As this factor has been instrumental in introducing a non-indigenous bias to education systems, there is an urgent need to encourage text-book writing from the perspective of the individual developing country. Although this is by no means an easy task, even the initial steps taken in this

[30] It has been estimated that the cost of the British Research Assessment Exercise expends 10% of available research funding on the process. See http://repp.anu.edu.au/Linkage%20grant.htm.

direction will encourage domestic scholars to look at social phenomena in fresh light and will represent a significant improvement over the current practice of simply transferring information from developed countries without any major attempts at modification. It would therefore be appropriate to attach much more importance to text book writing in evaluation schemes, especially if it involves team effort. Emphasis on text book writing would contribute to the development of a native academic language as well as a more relevant and indigenous approach to social science research which would in the final analysis lead to the emergence of a viable academic community.

Second, a dynamic domestic academic community[31] with close interaction among its members is what is lacking in most developing countries. In evaluation processes at all levels, but especially at the level of academic institutions much importance should be attached to this objective. On the basis of their success in developing such a community at the local and national levels individual developing countries can join the international academic community from a position of strength. To this end, there is an urgent need to establish a scheme encouraging the flow of scholars among different academic institutions over short periods. It is also necessary to integrate graduate students to academic life by increasing staff-student interaction through workshops, seminars, and conferences. To strengthen graduate programs, there is a need for universities in the same province or even region to increase their cooperation also by pooling their resources. Again to serve the objective of developing a domestic academic community it would be appropriate to have better dissemination of information among academics through the internet and publication of abstracts of papers and ongoing research.

The rapid growth in the number of universities in which tuition is in English should be checked and existing ones should be scrutinized with a view of confining this practice to several institutions at most. While accepting the importance of using course material written in foreign languages, one finds it totally absurd that students in these universities should learn subjects like say, the Turkish economy, Turkish economic history, and the Turkish legal system in a foreign language. What is even more absurd is the requirement by these universities that students should write their master's and doctoral theses in a foreign language.[32] The emergence of an academic community is important for yet another reason. The very existence of such a community will be instrumental in drawing the best minds to academic research and will automatically create pressure for increased research activity and may be instrumental in encouraging the laggards.

Third, the overemphasis on research output in appointments and promotions has led to the diversion of resources to spheres with questionable social relevance. As one observer has rightly pointed out "The fact *that* a scientist publishes, *how much* he publishes, and especially *where* he publishes, has become much more important than *what* he is publishing"[33] Making social science research more relevant and increase its practical use should therefore be another major tenet of the new approach to evaluation in social sciences. The big challenge therefore is to carry this issue beyond the individual level pertaining to academic appointments and promotions into the social realm and channel research effort into socially more productive areas. This challenge is probably the strongest in the field of economics which has perhaps moved faster and deeper in the wrong direction than other social science disciplines. The deep penetration of highly mathematical and technical approaches has been instrumental in the erosion of much of its social content to such an extent that even the inclusion of economics

[31] See Tekeli (2003) on the importance of this issue.
[32] As a partial solution to this problem, graduate theses in social sciences at the Middle East Technical University are now required to include an extensive summary in the Turkish language.
[33] See Holub, Tappeiner and Eberharter (1991: 326) as cited in Laband and Tollison (2003: 161).

within the social sciences has become questionable. On the other hand, economics is well placed to bring fresh thinking into this area not least because efficient allocation of resources and social cost benefit analysis are subjects that fall well within its domain.

Evaluation schemes should attach much more weight than at present to socially relevant education and research activity with a problem-solving focus. In the sphere of education, contributions towards making existing curricula more relevant to societal requirements and developing new courses with this objective in mind should be given increased weight in the evaluation of both institutions and individuals. On the research front too, there is an urgent need for reorientation of research activity to socially more relevant issues. For example, at a time of deep economic crisis in Turkey in recent years which was accompanied by a severe recession, galloping inflation, increasing unemployment and poverty, Turkish economic research activity was still geared to an obscure emphasis on international publications *per se*. The fact that only a very few of these publications were addressed to these pressing socio-economic problems remains an unresolved paradox of existing evaluation schemes. As a natural outcome of this lopsided incentive scheme, the initiative in directing the future course of the economy and drawing up its research agenda was to a large extent captured by international financial institutions and their specialists whose academic background and expertise in general lagged considerably behind some of their Turkish counterparts.

The emphasis on socially relevant research and domestic publications should not, however, mean a totally inward orientation in academic activity. Instead, the objective here is to redress the bias of existing evaluation schemes against indigenous research and publication efforts directed at indigenous problems. An important step in this direction would be to encourage comparative interdisciplinary research related to say, Turkey and another developing country which would broaden the scope of researchers, provide useful spillovers for their research capabilities, and improve the quality of their research. Comparative studies of this type would no doubt be a better and safer way of joining the international academic community which, as it stands, exhibits one way traffic-from developed to developing countries.

One should recognize, however, that defining what is relevant may be problematic especially in undemocratic developing economy contexts where this objective may fall victim to powerful interest groups, pulling domestic research agendas into conservatism. As a safeguard against this, a broader participation may be sought at various stages of research activity. A useful approach in this respect would be to increase the interaction between the various social groups, non-governmental organizations, and public administrators who are more familiar with the problem that needs to be solved and the academics at each stage of the research process. Research projects aimed at solving deep-seated socio-economic problems, engaging the considerable research potential that exists in the research departments of various public institutions as well as universities, and increasing the cooperation between different academic disciplines should be given priority. Apart from increasing the relevance of social science research, this approach may also contribute to the development of the domestic scientific community by increasing the interaction among scattered research groupings and increase their problem-solving ability.

It is encouraging to observe some signs of success in Turkey in all the three spheres identified above. For example, performance in education is included, albeit with much less weight than publications, especially international publications, amongst the end -of- year assessments of the performance of academic staff at the Middle East Technical University.[34]

[34] A foundation closely associated with the Middle East Technical University offers a small number of awards each year for excellence in teaching. Awards are based on student evaluations for individual courses at the end of each semester.

One of the must prestigious social science journals in Turkey, the METU Studies in Development, covered by some international indexes, is bilingual, publishing articles in both Turkish and English. Such intermediate categories may no doubt go some way towards bridging the gap between the purely local and international. The Higher Education Council of Turkey has recently started a program in which junior academic staff of universities in the periphery carries out their doctoral studies at more established universities and return to their previous universities upon completion of their studies. In the same vein, the Turkish Academy of Sciences started in 1996 a scholarship program in social sciences in which graduate students begin their studies in Turkey but are given scholarships to pursue their studies abroad for a period of up to eighteen months at the end of which they return back to Turkey to complete their studies. The selected programs at various universities are also given pecuniary support for each of their graduate students obtaining scholarships from this program to strengthen their own graduate programs.[35] This scheme can be regarded as efficient import-substitution to coin a phrase from economics, with such obvious benefits as giving the students experience in an international setting without separating them totally from the local setting and by requiring them to return to their home country, acting as an effective stopgap against brain drain.

While these are welcome steps in the right direction much more needs to be done to cope with the difficulties and trade-offs emanating from the current system of evaluation and make sufficient progress in the three spheres identified above. The evaluation of performance in social sciences is not simply a technical matter. It involves an exercise beyond quantitative analysis in the realm of a much thornier field, involving the measurement of quality. As one observer has perceptively argued, “the concept of quality is multidimensional. It is not at all obvious how to add up dimensions as scientific quality, productivity, scientific relevance, societal relevance and so on in one mark.”[36]

While accepting that research is a fundamental objective of universities that needs to be encouraged through appropriate incentives, we have demonstrated that the current practice of evaluation is problematic in a number of key spheres, especially in the context of developing countries. The latter call for a major rethinking and radical reform of the existing system of performance evaluation in Turkey. Before considering the question of what is to be done, it is perhaps pertinent to start with a much easier question of what should not be done. The first major step to be taken in this direction should be to concede that, despite some of its advantages the tendency to use international publications as the main anchor in performance evaluations has gone too far. There is an urgent need to remove international publications as a precondition for academic appointments and promotions and conduct the evaluation process at all levels independently of whether the research findings are published at home or abroad. Another measure aimed at reforming the existing system would be to evaluate performance on the basis of multiple criteria. Rather than simply counting the number of SSCI publications, a more flexible system can be designed giving a bigger weight to publications which have received a large number of citations. If international publications are to be emphasized, this should be done on a more selective basis, with theoretical contributions and comparative and multidisciplinary studies, for example given much higher weight than others. A precondition of increasing the relative importance of domestic publications in this way is that books and articles published domestically should be subject to objective and strict refereeing. Another step in the same direction should be to develop a comprehensive citation index for domestic publications and incorporate citations in it along with SSCI citations in

---

[35] See http://www.tuba.gov.tr for details of this program.

[36] See Hansen (1998: 73).

measuring impact. Similarly, international accreditation institutions should incorporate domestic publications into their evaluation criteria.

At the institutional level, the quality of research facilities that an academic unit can utilize, in particular the library not only in terms of the sheer size of its collection but also in the frequency with which the collection is updated, and the amount of resources devoted to subscription to academic journals can be included among important assessment criteria. Likewise, the number of research students and how well they are integrated to academic life in that unit should be taken into consideration. A criterion in this respect would be the level of cooperation in research and teaching among members of the same department, with members of other departments and even other universities at home and abroad. Assessments at the institutional level through self-assessments and/or by an external body[37] will also impose automatic pressure on individual researchers to improve their performance. Self-evaluations of academic units which are used in some countries such as the Netherlands as part of the evaluation process should perhaps be extended to the individual scholar, involving self-assessment over a certain period. Such information which may be conveyed through a short report will no doubt provide the assessors with some useful information about the obstacles faced by the slow performer which may extend well beyond the personal level to the level of the institution and indeed the whole academic community.

To improve the existing system of academic promotions, jury membership duties should be restricted to a maximum of three candidates per year, jury members and candidates should not be affiliated with the same academic institution, and evaluation should be based primarily on the three major publications as identified by the candidate. Reports of jury members should be written in the format of a book review or review article and should be published by the Association of Universities at the end of each round of assessments. Although this may be a further impingement on assessor time, there are important potential benefits by way of improving the quality of assessments. This would also increase the flow of information among members of the academic community and have other important spillover effects such as providing the candidates with expert opinion. The latter would be particularly beneficial as writing of book reviews, survey articles, and review articles which are an integral part of the process of implicit peer assessments in other countries are not much in vogue in Turkey.

The proponents of the existing evaluation schemes in Turkey based on research and publication performance should recognize the fact that there is an on-going debate on this issue in almost all the countries implementing them and consider the emerging criticisms with less self-righteousness. At the next step, they should consider the possible advantages of an alternative mechanism aimed at closer interaction between education and research, with emphasis on the development of a viable domestic academic community sensitive to domestic socio-economic issues over the current practice based heavily on the number of international publications.

[37] Rules and regulations which came into force in January 2003 with the aim of increasing quality of education and research in higher education envisage the internal evaluation of programs acording to centrally determined set of quidelines, putting particular emphasis on participation and transparency at all levels of the evaluation process. The Academic Evaluation and Control of Quality Committee set up for this purpose is assigned the task of oversseeing the internal review process of individual universities against the background of quality and efficiency objectives. See http://www.yok.gov.tr/uak/yonetmelikler/kalitekontrol.html for details. It is early days to say whether this will develop into an effective evaluation process.

REFERENCES

Berrington, H. (1998), "Research Assessment in Britain", in DISRR (1998), 11-14.

Brauninger M. and Haucap, J. (2003), "Reputation and Relevance of Economics Journals", *Kyklos*, 56 (2), 175-198

Butler, L. (2003), "Explaining Australia's increased share of ISI Publications- The Effects of a Funding Formula Based on Publication Counts", *Research Policy* 32, 143-155.

DISRR (The Danish Institute for Studies in Research and Research Policy) (1998) "Assessing the Assessments European Experiences", http://www.afsk.au.dk/ftp/Assessment/Assessment.pdf.

Fish, S. (2003),"The War on Higher Education", http://www.ephilosopher.com/article680.html.

Geuna, A and Martin, B.R (2001), "University Research Evaluation and Funding: An International Comparison". SPRU Electronic Working Paper Series No. 71, http://www.sussex.ac.uk/spru/publications/imprint/sewps/sewp71/sewp71.html.

Hansen, H.F. (1998), "Methodological Problems / Warnings Discussion of the Validity of the Recommended Criteria" in DISRR (1998), 67-75.

Holub, H.W., Tappeiner, G. and Eberharter, V. (1991), "The Iron Law of Important Articles", *Southern Economic Journal,* 58, 317-328.

Laband, D.L. and Tollison, R.D (2003), "Dry Holes in Economic Research", *Kyklos*, 56, (2), 161-174.

van Langenhove, L. (1998), "Evaluation of Social Sciences Rtd in Belgium", in DISRR (1998), 34-36.

Newton, K. (1998), "The Case of Danish Political Science seen from the UK", in DISRR, 15-18.

Siune, K. (1998), "Why Assess Assessments?" in DISRR, 5-7.

Şenses, F. (1986), "Başlıca Sorun Alanları Açısından Türkiye'de İktisat Eğitimine Toplu bir Bakış", *İktisat Dergisi*, 254, January, 7-18.

Şenses, F. (1993), "Akademik Yaşamımıza İlişkin Bazı Gözlem ve Öneriler", *ODTÜ Gelişme Dergisi*, 20 (1-2), 169-88.

Tekeli, İ. (2003), "Dünyada Bilimsel Faaliyetlerin Kurumsallaşmasında Yaşanan Dönüşümün Sosyal Bilim İnsanlarının Performanslarının Değerlendirilmesinde Getirdiği Değişimlerin İrdelenmesi, Sosyal Bilimlerde Performansın Değerlendirilmesi", Atelye Çalışması'na sunulan tebliğ, 26-27 Haziran 2003, Orta Doğu Teknik Üniversitesi, Ankara, mimeo

Yurtsever, E., Gülgöz, S., Yedekçioğlu, Ö.A and Tonta, M. (2001), "Sosyal Bilimler Atıf Dizininde (SSCI) Türkiye", Türkiye Bilimler Akademisi, Ankara.