CHANGES IN RESEARCH ASSESSMENT PRACTICES IN OTHER COUNTRIES SINCE 1999:

FINAL REPORT

28 FEBRUARY 2003

Authors: N von Tunzelmann, E Kraemer Mbula

INTRODUCTION

As requested by HEFCE, we are reporting on key changes in research assessment practices in a selected group of countries where we have prior evidence of substantive change since 1999, when SPRU's previous report (Geuna et al. 1999) surveyed the general picture. Some notes towards the present report were despatched to HEFCE in December 2002 and an interim report supplied on 7 February 2003. We now present the final report for this contract. Within the time and budget available, we have not been able to produce an exhaustive account of all changes taking place in all countries regarding research assessment, but we feel that the material contained in this report gives a fair picture of the extent and direction of change that has taken place over these three and a half years since the 1999 Report.

For the present report we were also requested to obtain material where possible on several new issues currently on the UK RAE agenda; an overview of what we have found appears in the summary below, and details are given in relation to specific countries thereafter.

For our initial submission in December we relied almost entirely on web searches and available documentation. In the interim report in early February we were able to extend these and also supply a moderate amount of information from correspondents abroad. For this final report we have been able to widen the net of response to a considerable degree. However some promised material has not arrived in time for submission of this report; should it do so within a reasonable time-frame and prove to alter our findings in any substantive way we will prepare an addendum as and when.

SUMMARY OVERVIEW

The 1999 Report developed a methodology for considering various types of research assessment. Specifically, the results for individual countries were grouped into:

1) Evaluation performer (national level and institutional level) and evaluation purposes (summative or formative);

2) Evaluation criteria (quality, impact, etc.) and methods;

3) Evaluative remarks (strengths and weaknesses, impact on research and teaching).

Most attention was then paid to the first two points. For this present report on the changes since 1999 we have found few instances of change in regard to evaluation performer under point (1). We do however get a sense of a switch of emphasis from summative towards formative purposes of evaluation; or more precisely, perhaps, it could be claimed that countries embarking on more formal research assessment exercises are giving greater weight to formative functions than they envisage as existing in the UK RAE (which is often cited as a 'benchmark'). Examples below include Ireland, Flanders and New Zealand. The chief justification for more formative measures appears to be giving a greater degree of individual 'ownership' of performance improvement, e.g. for making more distinctive choices of a strategy for achieving improvement.

A possible exception is the case of Taiwan, where however a rather different situation seems to apply – Taiwanese studies allege that the country's industry is very strong by international standards, e.g. in patenting, but its academic publication record is a lot weaker (e.g. Luo, 2002), i.e. arguably the reverse of the UK situation. The great merit of the UK RAE is seen in Taiwan as its boosting of the competitive spirit in universities (ibid.). We have received some rather contradictory messages from correspondents in Taiwan about what is actually happening there, and our judgements may thus be incorrect.

As in the 1999 Report, we cannot claim to have a great deal of evidence to bring forward regarding point (3), evaluative remarks. Most countries still regard their systems as somewhat experimental. The most obvious kind of inference we can make is that the nature of changes

introduced is a likely reflection of dissatisfaction about previous systems (or indeed the previous absence of systems).

Our discussion in the present report therefore has more to say concerning point (2), the evaluation criteria and methods. A partial critique of existing methods was given in the 1999 Report. The UK RAE is frequently criticised for not using a broader basis of bibliometric indicators (our own colleagues have been known to voice these views, amongst many others). While the objections to such broadening of the indicators in the UK are often made on what to us look somewhat specious grounds, we have to admit that the only case we know of in which a more thoroughgoing adoption of formal bibliometric criteria has been made, namely Flanders in Belgium, does raise serious doubts about the practicality of implementing such methods in a country like the UK. Again, the experiment is however too recent to be properly assessed.

In the event, our present work is rather directed to an issue which falls under several of the above headings, which can be referred to as *evaluation governance*. In line with standard definitions of 'governance' (as comprehending structure, control and process), this includes the structures of decision-making (which relates to point (1) on performers) but also processes (which relates more to point (2) on methods).

The 1999 Report drew attention to there being 4 categories of countries in regard to university research funding practices. The first used a performance-based approach like that in the UK, using research evaluations to distribute funds at least in part. Poland, Slovakia, Hong Kong and Australia were included in this category. The second category comprised countries that used another indicator than research evaluation (though sometimes with a small portion of this), generally student numbers or similar. This was the largest group in 1999 Report (Germany, Italy, the Nordic countries, Hungary, New Zealand). Third was a small group of two countries (Austria and France) where funding allocations were described as open to 'negotiation'. Fourth was a group of 3, the Netherlands and the USA and Canada, where research assessment and funding were separated.

The countries included in the present study do not fully overlap with those examined in 1999 (for example we have not looked at Eastern Europe this time), but it seems reasonable to assert that there has been some shift towards the first category. The Netherlands and New Zealand have both made explicit moves in this direction, though in both cases during the

course of our enquiry (January 2003 and December 2002 respectively). Other countries like Taiwan, not considered in 1999, seem to be moving in the same direction. On the other hand, the North American countries have not really budged insofar as university research is concerned, and Italy too has seemingly not changed.

Some of the other key conclusions from our study are the following.

- A general conclusion reached is that countries like Ireland, France, Switzerland, Denmark, Japan and New Zealand are intending to rely to a greater degree on self-evaluation, subject to oversight panels (the USA does the same in its benchmarking of Federal research agencies). This is intended mainly as a way of sharing the burden between centralised and decentralised agencies, and perhaps giving a more rounded picture. It is not transparently clear how much this would differ from the submissions made by universities as the first stage of the existing UK RAE, but certainly the countries concerned appear to feel it is indeed different. It is in this way that we draw the inference of a shift to a more formative role for research assessment in such countries. All countries introducing such schemes appear to be aware of the probable 'puffery' of self-evaluations, which the oversight panels are intended to bear down upon, but the same could be said for the existing UK system. Self-evaluations may need to be linked to more objective measures of performance benchmarking if they are to have any real standing (see below). In favour of self-evaluation is said to be the opportunity for making special anti-discriminatory cases e.g. in relation to the treatment of women (see below).

- While there is widespread interest in the UK RAE, intensive evaluation schemes are mainly limited to smaller countries or regions, and indeed some of our correspondents have stated to us that they would be infeasible in larger countries. This suggests that the UK has problems of intensive application in a comparatively large country, as of course has often been expressed via the resource cost involved. This perhaps suggests that the practicability of a more devolved system (possibly associated with the recent emphasis on RDAs) might be investigated. Taking this one stage further, we might suggest coupling 'intensive' evaluation with the previous point about self-evaluation, i.e. getting universities to produce their own benchmarking using bibliometric or other methods.

- We would also draw attention to the basic theoretical underpinnings of any research assessment, as noted in relation to Ireland. This overlaps with concerns expressed in some

quarters that the RAE is too narrowly academic, though it goes rather beyond that to deeper issues of the 'contribution of basic research to the modern economy'.

- Finally we could mention that most of the evaluations occur with longer gaps than has been the past practice of the UK RAE, though the latter is of course now changing.

We were also asked to report on any evidence located with respect to four supplementary issues (again we have been unable to provide an exhaustive analysis):

a) Use of benchmarking practices:

This is attracting growing interest though mostly independently of university research assessment (see evidence below for Ireland and the USA), with the US adoption of the 'virtual congress' worthy of note. However 'benchmarking' covers many different approaches at many different levels, and a more concerted study would be needed to draw any strong conclusions. We make the point above that, if benchmarking is to be adopted for assessing individual universities and colleges, it might be necessary as a practical matter to devolve the responsibility to them for furnishing the data, since all known benchmarks suffer from a variety of limitations (see the discussion of Flanders). There is also some concern about over-use of benchmarking as encouraging 'moral hazard', e.g. undue focus on attaining the specific benchmarked targets.

b) Special treatment of the social sciences and humanities:

We found relatively little reference to this in research evaluations, though some countries have weighting systems (see evidence for Ireland, Flanders, Japan and New Zealand, also references to Hong Kong and Australia), while many express general concerns about perceived imbalances in the educational system or student uptake.

c) Special reference to the treatment of women:

This area in a general sense has been recently overviewed by the European Commission, particularly through the 'Helsinki Group on women and science' (1999), which has produced national reports on the situation of women scientists in EU countries and represents an important base to compare different experiences across countries. From these and other sources we include some remarks on individual countries. What we found less on were the specific allegations of implicit discrimination through the use of RAEs themselves, although it does appear to be widely accepted that this occurs (see the note for New Zealand).

d) Explicit use of the UK RAE model and the UK debate:

As stated above there is clearly near-worldwide interest in the UK model, which itself has become a 'benchmark' for research evaluation of higher education. However it has not yet been precisely replicated elsewhere, and most countries are probably holding a 'watching brief' (see the evidence for Ireland, Australia and New Zealand). We have learnt recently of a proposed implementation of a UK-style RAE in Taiwan, although as implied above we are getting somewhat mixed messages about how far this has progressed, and indeed what its precise form will be.

COUNTRY STUDIES

In this present report we follow the order of countries used in the 1999 report, namely Europe, Asia, Australasia, North America; though as stated above the selection of countries is somewhat different.

IRELAND

Ireland has gone through dramatic changes in public policy towards Science, Technology and Innovation in recent years. The launch of the *National Development Plan 2000-2006* represents a high commitment of the Irish Government to scientific and research activities.

A reason for these changes can be found in the observation of the fact that, although the public funding for ST&I has approached international norms, Ireland still lags behind OECD countries in terms of evaluating the benefits of such funding provision. A scarcity of expertise has been identified in terms of evaluating the outcomes of expenditure on ST&I within the Irish public sector bodies.

ICSTI (the Irish Council for Science, Technology and Innovation) has recently published a number of reports, including the report of a task force on Embark Initiative 2002 'Measuring and Evaluating Research'. This report looks at the multiple evaluation practices in different organizations relating to STI. Examples of units with evaluation units include: the Higher Education Authority (collection of all personnel data, ad hoc evaluations of all programmes), the Universities with a quality review system (QAQI), Teagasc (for agriculture, food and rural development), and Enterprise Ireland (evaluating all their technological programmes). The Council concludes that inputs are quite well monitored in Irish evaluations, but not outputs or impacts.

To account for this, the report begins by contrasting the traditional 'linear model' approach, also implied in neoclassical economics approaches, with the broader 'evolutionary/ institutional model' operating typically through 'systems of innovation', allowing for multiple feedbacks and interactions. In coming to measurement, the former leads to an 'input-output' approach whereas the latter leads more naturally to a 'throughput' approach in which the focus lies more on the process of research rather than the products. Frequently, though not

necessarily, the former measures are used in ‘summative’ fashion to inform a particular policy decision (like an RAE rating) while the latter are used in ‘formative’ fashion to improve performance.

Principal indicators and evaluation methods in use tend to be implicitly related to the former ‘input-output’ approach, at various levels from the macro to the micro – these include: R&D expenditures, human capital indicators, bibliometric analyses, patent analyses, technological intensities of production, technological trade indicators (technology balance of payments, technology levels of exports), growth accounting analyses, measures of price and quality changes.

Techniques suggested that might be used in ‘throughput’ analyses include: peer review, surveys/interviews, case studies, cost-benefit analyses, productivity spillover models, knowledge flow models (linkages), technology foresight exercises. However it might be interjected that the ways in which these are generally implemented fall some way short of theoretical desiderata, and the resource cost of doing them well could be extremely high. On the other hand, the ‘throughput’ approach is much more in line with modern thinking about good innovation practice.

In this vein, the report draws in part on the work in the UK at SPRU, while other Irish reports also refer to work from PREST.

Altogether there are 13 funding agencies identified in Ireland (Embark Initiative 2002). In terms of subject breakdown, the Research Councils in Ireland are divided into Humanities and the Social Sciences (IRCHSS), Science, Engineering and Technology (IRCSET), and Health (HRB), plus some smaller institutes. The first has recently launched a Project Funding Scheme to support team-based research in relation to economic, social and cultural development. The second launched its first programme, the Basic Research Grant Scheme, jointly with Enterprise Ireland, in December 2001. Science Foundation Ireland was founded as a result of a 1998 Foresight Exercise in 2000 and currently aims at recruiting and retaining research groups and centres, with biotechnology and IT identified as the main targets. Much emphasis is placed on international as well as national peer review, and on European projects.

Regarding quality of research, the University Act of 1997 required each Irish university to review the quality of their research work on a ten-year basis. This quality review system is called the ‘QAQI Programme’ (Quality Assurance/Quality Improvement) and is managed by

the Quality Assurance Office. The programme is based on self-assessment of the unit mechanisms and peer-review by external agents leading to a 'Quality Improvement Plan'. The introduction of the quality standards within the UK RAE system in Britain in the eighties had a profound effect on the subsequent development of QAQI procedures in Ireland. The importance of quality in Irish Universities is clearly underlined by the funding (€1.5M in 2001 and €2M in 2002) that has been made available by the HEA under the National Development Plan.

As a collaborative measure, the governing authorities of the seven universities involved in the QAQI Programme have established in 2002 the '*Irish Universities Quality Board*' (IUQB), which plays an important role in the selection of the agencies responsible for the periodic reviews and evaluation reports.

a) use of benchmarking:

Benchmarking exercises have not been systematically repeated in Ireland, but some examples can be found within the last 4 years:

- ICSTI carried out a benchmarking exercise on science, technology and mathematics education in 1999. Its report '*Benchmarking School Science, Technology and Mathematics Education in Ireland Against International Good Practice'* was released in February 2000. Both qualitative and quantitative indicators were used for this study and many weaknesses of the Irish Education system were brought together in comparison to the benchmark selected countries (Scotland, Finland, Malaysia and New Zealand).

- A more recent benchmarking exercise was conducted by ICSTI whose final report called '*Benchmarking Mechanisms and Strategies to Attract Researchers to Ireland'* was released in February 2001. The UK, US, Netherlands, Finland and Denmark were the five selected benchmark countries[1]. Some recommendations for Ireland derived from this exercise are: (1) Build up centres of excellence, (2) Improve international networks and visibility of Irish universities, (3) Improve the status and remuneration of research graduates and post-doctorates, (4) Make the move to Ireland as smooth as possible.

b) treatment of natural/social sciences:

Regarding the specific consideration of natural sciences against humanities/social sciences, this issue is sometimes linked to the gender issue. The main actions taken have been over student intakes into physics, and to girls in physical sciences in secondary schooling.

c) treatment of women:

The national report on the situation of women in Ireland produced by the Helsinki Group (July 2001) identifies three key Government Departments taking positive actions to increase the participation of women in mainstream science, technology and research in Ireland. These are the Departments of (i) Education & Science (DES), (ii) Justice, Equality & Law Reform (DJELR), and (iii) Enterprise, Trade and Employment (DETE), and their agencies. In this report it is stated that "The National Development Plan 2000-2006 (NDP) aims to promote equal opportunities between women and men through the strategy of gender mainstreaming". This promotion is reflected in several actions. The most recent related to research activity are summarised in the following table:

DEPARTMENT	*INITIATIVE*	*FUNCTIONS*
DES	• Establishment of the Equality Unit	Functions: co-ordinate and monitor the process of mainstreaming a gender perspective into all areas of the educational system.
(HEA)	• Establishment of the Higher Education Equality Unit	Functions: promote and encourage good policy and practice in relation to the tackling of inequalities in the higher education sector in Ireland.

NETHERLANDS

The situation in the Netherlands was described at length in our 1999 Report (pp. 25/33). In that Report we looked at the 'Protocol 1994 for Quality Assessment of Research' (p. 27), which was a uniform protocol for assessment of basic research of all disciplinary areas. This protocol was updated in 1998 in the 'Assessment of Research Quality 1998', valid until 2003. According to our contact in Netherlands, Don F. Westerheijden, "its assessments were

[1] UK and US as examples of top global competitors in attracting internationally mobile researchers and the Netherlands, Finland and Denmark as "competitor" European countries with similar small research systems.

generated for internal use in the universities only and, in opposition to the UK RAE, there was no connection between the evaluation and the level of received funds. It is worth mentioning that, simultaneously, other evaluation exercises were taking place at the national 'research schools' by the Royal Academy (KNAW) every five years, and the council (NOW) also had its project evaluations for funding purposes."

Not until very recently (January 2003) did the VSNU together with NWO and KNAW publish a new protocol to consolidate most of these evaluations into a single new procedure. This document is named '*Standard Evaluation Protocol 2003-2009 for Public Research Organizations*' and is derived from the 2001 report 'Kwaliteit Verplicht' (Quality Obligations) of the group for Quality Assurance in Scientific Research. The main objectives of this protocol are to be both improvement and public accountability of research activity. In relation to the old protocol, this new one puts greater emphasis on prospective analysis and extends the aspects of quality of research to wider socio-economic impacts of research.

According to this document "the [new] system aims at operating with the least possible burden on researchers: a self-evaluation once every three years and external evaluation every six years." It also encourages the accessibility through the Internet of certain relevant data by the generation of a national research information system, although this facility was not yet available by the end of 2002.

In line with other European countries that will be examined below (also Japan and New Zealand), with the new procedure the Netherlands is giving a more central role to self-evaluation. The evaluation process is composed of nine steps, as follows:

1) Planning and timetable for all research institutes, including a protocol draft for each specific evaluation
2) Protocol for the specific external evaluation
3) Selection of the evaluation committee
4) Self-evaluation
5) Evaluation committee's working programme
6) Evaluation Report, including a review of the entire institute and a review of each research programme
7) Conclusions by the Board, based on the self-evaluation document and the final evaluation report
8) Making the evaluation results public

9) Public meta-evaluation, carried out by an independent committee

Each institute needs to be assessed by an external peer evaluation committee once every six years and undertake a self-evaluation every three years.

The protocol for the External Assessment of Educational Programmes 2000-2005, entitled '*Quality Assessment Made to Measure*', was also released recently, in July 1999. This protocol does not attempt to be prescriptive but it is suggested by the VSNU as a guide in order to ensure that "educational assessments are carried out in an atmosphere of open discourse, proceed efficiently and [...] are effective". This protocol is the basis for the third round of university study programme assessments by a review committee and it "broadly follows the approach used in the previous two rounds (1988-1993 and 1994-1999) but also departs from it quite significantly in some areas."

The reason for these changes can be explained by an adaptation of the evaluation system to more general changes in the university world in recent years. According to this document, Dutch universities have gained autonomy in the design of their own organizational parameters and in the formulation of their own missions at the same time as they have been receiving more funding from external sources. In this sense, the main changes from the last rounds are related to:

- Greater attention to the specific characteristics of the discipline, the institution and the study programme ('made to measure')
- The opportunity to ask the review committee for targeted, confidential advice (the 'management letter')
- A great emphasis on evaluating the study programmes in an international context.

a) use of benchmarking:

It can be seen that the international dimension of the educational and research assessment has gained great importance within the last few years in the Netherlands. We are aware that the Dutch government has been active in benchmarking the whole economy including higher education and science, and currently takes part in the EU's open coordination mechanism; nevertheless there is no evidence of any systematic benchmarking exercises for research. However, some benchmarking exercises have been performed for specific branches of sciences that have followed the VSNU protocol. These exercises can be found in the specific

reports (rapports) at the VSNU website from 1999 to 2002 in a number of fields such as Economics and Aerospace Engineering.

b) treatment of natural/social sciences:

From the new protocol and the impressions of our contacts in Netherlands, it can be said that the aim is to standardize the evaluation methodology across the disciplines, covering the humanities/social sciences as well as the natural sciences. Nevertheless, the new protocol is found to be primarily directed toward the evaluation of scientific research. For that reason, the extended use of quantitative indicators for the measurement of productivity (scientific output), which have gained credibility in the physical and life sciences, still remain problematic in the social sciences and humanities where publication patterns may vary widely between disciplines. In order to avoid this problem, the protocol promises the consideration and close observation of the development of "new tools into the evaluation process once they have proven their credibility and can provide significant value to the evaluation process."

c) treatment of women

General Dutch councils have taken gender issues into account. The Advisory Council for Science and Technology Policy has published advice on Women in Science in January 2000 and the Education Council advised on gender policy in education in October 2000.

According to the Helsinki Group's report (May 2002), "Since the beginning of 1999, NWO, VSNU and the Ministry of Education, Culture and Science have developed a national programme to promote more women in associate professor positions. After the first call of this ASPASIA-programme 68 female researchers were promoted to associate professor. Among them 30 received an NWO-grant for a research project to be carried out by a PhD. The second call is held in 2002. The programme is successful, and the quality of applicants is high. The Parliament has supported the programme by voting an extra €800,000 in November 2000. The total budget is about €7.5 m."

d) influence of UK RAE:

According to our contact expert in the Netherlands, David Campbell[2], "the UK RAEs were extremely influential in starting the comprehensive ex-post university research evaluation in the Netherlands. Taking advantage of the fact that the Netherlands represents a medium-sized Western European country, the Dutch evaluation system operates in a more disaggregated way than in the UK."

FLANDERS (BELGIUM)

The Flanders region of Belgium is the only region we know of that has made a concerted attempt to introduce bibliometric methods into evaluating university research performance. Six universities are involved, ranging from small to medium/large in size. FWO-Vlaanderen, the Fund for Scientific Research in Flanders, monitors a large portfolio of basic research grants and projects to individual researchers (including PhD students and post-doctoral grants) and academic promoters at Flemish universities. The selection and monitoring mechanism is conducted by scientific commissions that base their decisions on a peer-review system, consistently involving foreign experts in evaluating the proposals submitted to the agency.

Besides the public R&D funding via FWO-Vlaanderen, which is distributed on a project-per-project basis or on an individual basis, the Flemish government created a mechanism that allows for supporting more large-scale basic research at universities. Except for setting certain quality guidelines and performance expectations, the government does not intervene at all in the *internal* selection and monitoring process for the grants. The mechanism thus created has been called 'Bijzonder Onderzoeksfonds' (BOF) and had a total budget of €90 million to distribute across the 6 Flemish universities for fiscal year 2002. The weights were at first based purely on student numbers, according to a weighted scale.

The dissatisfaction with a numbers-led weighting is now being corrected by introducing explicit bibliometric indicators. This has now led to the creation of a dedicated research and policy support staff, called 'Steunpunt O&O Statistieken' (SOO), to support a major inter-university funding allocation decisions (Debackere & Gänzler, 2002). Bibliometric data have

[2] David Campbell will soon release a four-country comparison of university research evaluation: David F.J. Campbell (2003), 'The Evaluation of University Research in the United Kingdom, the Netherlands, Germany and Austria', in: Philip Shapira and Stefan Kuhlmann (eds.), *Learning from Science and Technology Policy Evaluation*, Cheltenham: Edward Elgar (forthcoming).

for the first time been used to allocate €93 m of public research money between these 6 Flemish universities for the fiscal year 2003, based on Web-of-Science SCI data (also SSCI and A&HCI) provided to SOO via a licence agreement with Thomson-ISI. Although the use of WoS data for evaluative and distributive purposes is not without controversy, they were considered the "best available, recurrently accessible, transparent and controllable" data for such a purpose (in SPRU we perhaps take a more critical view of WoS). In addition patent data (USPTO, EPO and WIPO) and innovation data (CIS) are collected. Older criteria such as student numbers retain a 50% weighting in the overall allocation formula.

Though our correspondent (Debackere) is one of the architects of this system, he is partially critical of what it is currently trying to achieve. Many assumptions had to be made in preparing the data for comparison. Many of these are well recognised in bibliometrics work, such as problems of co-authorship across institutions, lead authorship, fractional authorship, self-citation, etc. Problems of mis-spellings and different listings of individuals and affiliations all had to be cleaned. The magnitude of this task was what led our correspondent to note that he could not see the feasibility of doing anything similoar for a large region/country such as the UK.

Even then there are limitations on what is being achieved. Because of the particular limitations of SSCI and A&HCI, it is not being used for allocations in the social sciences or arts and humanities. Moreover the issue of 'impact' has not yet been fully addressed, as distinct from numbers of citations, Finally, our correspondent would wish the scheme to be extended to intra-university funding. He is however deeply concerned about over-use of any such tools, as likely to divert activity away from more academically useful research into tactics for cultivating citations.

c) <u>treatment of women</u>

The Helsinki Group's report for Flanders highlights the increasing importance that the gender issue has achieved in general terms in Flanders since 2000. The Flemish Minister for Equal Opportunities is authorised to pursue horizontal and vertical equal opportunities policies. In 2000 44% (€1.26 million) of the total budget for equal opportunities policy was spent on "specific actions for women".

The under-representation of women in science and research has become more visible since 2000 with the publication of the ETAN (European Technology Assessment Network) report and developed in a series of actions. In April 2000, a Flemish steering group for women and science issues, with broad membership, had its first meeting. In October 2000 an Inter-university Workgroup on Equal Opportunities was established within the Flemish Inter-university Council, charged with formulating recommendations on actions and measures to promote equal opportunities for men and women in Flemish universities. Since 2001 the *Support Centre for Equal Opportunities*, a consortium of research groups positioned between the universities, the administration and policy-making, is focused on the target group of women within the policy-oriented research field, with a budget of about €4.88 million from 2001 to 2006.

FRANCE

Our previous report explained the situation in France (pp. 33/6). It was observed that "no evaluation mechanisms were specifically designed to evaluate university research". In response to this lack of agreement on evaluation techniques, the '*Comité National d'Evaluation*' (CNE) launched in January 2001 the '*Guide d'Evaluation des Universités*', in which the procedures of evaluation are detailed. This evaluation is conducted by the CNE and has two phases: one internal and one external.

The internal evaluation is organised by the institution. This evaluation is helped by the 'guidelines for evaluation' and must involve all the institution's staff. It consists of the preparation of a report of internal evaluation that will be the main guideline for the external evaluation.

Second, the CNE organises and coordinates an external evaluation based on a peer review. The experts include university professors, higher education administrative or technical senior executives and key economic professionals, be they French or not.

The final evaluation report is prepared by the CNE on the basis of the three main sources of information: the self-evaluation report, the external peer-review report, and the visits organised by the CNE. This final report has to be confirmed by the president of the evaluated institution.

According to our contact expert in France, Bastiaan de Laat, "… there has been an increasing emphasis on evaluation of research during the last few years. This fact has not been necessarily reflected in the use of new evaluation techniques in general." Nevertheless, impact assessment and other alternative mechanisms of evaluation have been already introduced in certain research institutions. More research should be devoted to the case of France to clarify these particular aspects.

SWITZERLAND

The Switzerland Science and Technology Council (SSTC) is the advisory body of the Federal Council for all matters relating to science, education, research and technology policy. SSTC formulates general concepts for the attention of the Federal Council and suggests measures for their implementation.

During 2002, the SSTC conducted an evaluation of the Swiss National Science Foundation (SNF) and the Commission for Technology and Innovation (CTI). The report was delivered to the Federal Council on May 31st 2002. The focus is to be on the role of the SNF and the CTI in the Swiss system of promoting research, development, technology and innovation. The results of this evaluation should help in establishing the governmental goals for education, science, and technology policy 2004-2007.

The evaluation of the SNF and the CTI is a three-step process:

- Self-evaluation (including SNF and CTI basic reports)
- External experts' evaluation (including an expert report).
- SSTC's final report, which includes recommendations to the Federal Council. Its report is based on the two basic reports (self-evaluation and expert report).

In 1998, when the SSTC reviewed the existing programmes to prioritise research investment (Swiss Priority Programmes - SPP), they found that they were effective in triggering university-industry cooperation and interdisciplinary research, but not in leading to enduring centres of excellence. Following from that observation, in 2001 Switzerland inaugurated the 'National Centres of Competence in Research' (NCCR). This was initially composed of 14 new projects in five fields (i.e. life sciences, humanities, environment and sustainable

development, information technologies, and others). These projects promoted the establishment of research networks (e.g. groups of researchers from the academic, corporate and/or government sectors) around a core institution (mainly universities), which play a driving role. The following aspects are decisive for the approval of a Centre of Competence: it must conduct research of outstanding, internationally recognised quality, and actively foster knowledge and technology transfer, training, and the promotion of women researchers. A further aim of NCCR is in globally restructuring and improving the organisation of Swiss research.

According to Stefan Bachmann, our contact expert in the Swiss National Foundation, "the first round of site visits by international experts in all 14 NCCRs in the course of 2002 showed that the Centre's activity started well, with highly active management and [that] – besides high quality research – several structural effects on the Swiss university system are already visible". By its end the programme should include up to 25 NCCRs.

b) natural vs. social sciences:

Switzerland has been extensively debating an upcoming reform of the Swiss university system. In 2002, SSTC launched the report '*Structural Reform of the Swiss Higher Education System*'. The SSTC is highly committed to the enhancement of framework conditions for research in Switzerland. It has created different working groups corresponding to the following priority issues: Technology and Innovation, Humanities and Social Sciences, Clinical Research, and Career Development. These working groups are to elaborate SSTC's positions for the formulation of the government's objectives in terms of research policy, and for the upcoming debates on governmental goals for education, science, and technology 2004-2007.

Within Swiss universities, the imbalance in the humanities and social sciences between student numbers and faculty is constantly deteriorating, and fewer graduates are going on to post-graduate programmes of study. Targeted promotion measures are planned to be implemented (e.g. tenure-track assistant professorships), given that these forms of academic career development promotion have proven their worth in the Anglo-Saxon countries.

c) treatment of women:

In Switzerland in recent years, a multitude of organised mentoring projects for women have been initiated. The mentoring projects are running in the Swiss universities and research

institutions within the framework of the Federal Programme for Equality. Acting on a mandate issued by the Swiss Federal Government, the Swiss National Science Foundation (SNF) supports research undertaken inside and outside universities. The SNF has taken several initiatives for the equality and promotion of women in the Swiss research framework[3].

At the beginning of 2001, the SNF entrusted a '*groupe de réflexion*', composed of both internal and external experts, with the analysis of gender-relevant problems in SNF research support and the presentation of recommendations for any measures to be taken in this respect. Responsibility for a decision on further action lies with the National Research Council. In July 2001 the SNF administrative body established the position of Equal Opportunities Officer as a first step to implementing the recommendations of the *groupe de réflexion*. Moreover, it was decided to constitute a new, internal Equal Opportunities Commission in SNF as an accompanying committee for the implementation process.

DENMARK

The Danish government claims to be strongly committed to a significant improvement in research standards and reforming the structure of the Danish education system to support the research system.

The Danish Centre for Quality Assurance and Evaluation of Higher Education (EVA) was established in 1992 originally to evaluate education programmes in universities, but in 2002 it has undertaken complete institutional evaluations. The EVA develops and updates methods for evaluation across the entire educational system. These methods vary and are adapted to the various educational areas and levels.

Most of the evaluations are carried out according to a methodological concept consisting of three mechanisms:

- Self-evaluation
- Expert panels (peer review)
- Client assessments involving students, graduates and employers

[3] Collaborative work on this subject between SNF and SPRU has been proposed.

During the last few years the Danish government has been demanding more from research institutions in order to maintain the quality standards of Danish universities. New reforms are on board. According to our contact in Denmark, Finn Hanson, "A very significant change in research evaluation will take place in 2003 in Denmark. The Ministry for Science, Technology and Innovation has decided to use a benchmarking approach to evaluate the Danish universities on research performance (and education) with the help of the OECD. A pilot study on experiences and statistical information in some mid-size European countries (Finland, Austria) has just been finished and a conference introducing the new benchmarking approach was originally scheduled for late January but has been postponed to late February."

New methods of evaluation from the EVA and open competition for research funding are seen as other ways of raising quality standards. These pressures are seen as means to even greater co-operation between universities and the private sector.

In 2000 bilateral "development contracts" were introduced. The universities negotiate with the Ministry for contracts that specify the performance goals for the universities to achieve in both teaching and research. Formally contracts are a voluntary arrangement, but all universities have chosen to participate. The basis for negotiations is a proposal from the university on what should be its goals and ambitions. The contract period is four years, but the contract can be adjusted on a yearly basis. It is not a legal document, but it is seen by the Ministry as an instrument for dialogue and control (Smeby et al., 2003).

In this respect, Finn Hanson also asserts that "… a number of universities [have] set up local programmes to produce a better classification of the research production in the institution to prepare for the demands coming from the new contract steering system between ministry and universities about demonstration of productivity and strategies in research."

b) natural vs. social sciences:

Regarding the different treatment of social sciences within research evaluation, we note from our contacts in Denmark that the Danish Institute for Studies in Research and Research Policy (AFSK) initiated in spring 2002 a pilot study of benchmarking the research products (books, articles, etc.) from a number of social science institutes with a base in economics or business economics. Hanson points out that "the project is focused on using a special model for analyzing benchmarking (DEA) and has the ambition of including international comparisons. The very traditional approach to classifying articles and to the types of data to

be included has slowed down the project and up to now only a working paper on the ranking problem has been produced" (Langberg 2002).

c) treatment of women:

The question of gender in academia in Denmark has been discussed with a background of a number of studies showing a much slower change in the male/female ratio in university positions compared to other Nordic countries[4]. The increasing focus that has been given to gender issues within the field of research in Denmark is represented by the FREJA programme (**F**emale **Re**searchers in **J**oint **A**ction) launched in 1998. According to the last Helsinki Group's report, "The aim of the FREJA programme is to get more good female researchers into Danish research, to make them more visible, and to ensure more role models for female university students. Hence the projects were to be given to small groups instead of individual researchers … from all areas of research"; "… the Danish Research Councils who distributed the FREJA grants could give priority to a female applicant if two applicants were equally qualified. Moreover, the FREJA grant was given to a senior researcher who then received sufficient funding for a research group related to the project." The FREJA projects are presently running. They have not yet been evaluated but the impact of this programme has been very large within the Danish research community[5].

d) influence of UK RAE:

The influence of the UK RAE in Demark can be looked at in two ways. The first influence is the close combination of the outcomes of the evaluation and budgeting. In this sense, Hanson alleges that: "this is not yet on the policy agenda in Denmark even though there are political signals for a closer relation between some kind of results and budgets. Nevertheless, the contract steering system (it is only in its first 4 years) combines some results with budget issues and will be the model for Denmark over the next few years."

"The second influence of the UK RAE has to do with the problems detected in former RAE approaches, like quantity regardless of quality and the large number of resources used in the RAE evaluation. In this vein, there is a large number of articles and books (in Danish) on the more negative experiences of the UK RAE seen from outside."

[4] A study was published by the Ministry of Education in 1999 (B. Ståhle, Alder, køn og rekruttering i dansk universitetsforskning).
[5] A special research programme FREJA on recruitment of researchers with special focus on women will be reported in 2003.

More research should be devoted to the Danish case in order to specify the recent reforms in research evaluation, given the existing contradiction between Denmark's leading position as an advanced nation within the European context and the challenges it faces in research.

NORWAY

The Norwegian case was treated in our previous report (pp. 41/2). This information has been largely extended thanks to our expert contacts in Norway, Svein Kyvik and Jens-Christian Smeby. To stimulate universities and colleges to improve quality and efficiency a new funding model for basic funding of higher education institutions was introduced in 2002. The old model was mainly based on historical traditions and the number of students. The new model is to a greater extent based on performance in teaching and research. It differentiates between a performance-based grant for teaching, a performance and strategic based grant for research, and a basic grant. The performance-based grant for teaching is first of all allocated on the basis of earned credit points in the respective fields. The criteria for performance-based research funding are among other things based on the number of associate and full professors and grants from EU and the Research Council of Norway.

So far scientific publishing is not included in the formula because of a lack of reliable data. It is currently being discussed how a database including numbers and types of publication should be established. The distribution of the basic grant is mainly based on the number of students, infrastructure, and regional and national priorities in research and education. The different elements in the model were constructed in a way that did not change the allocation of resources between institutions in the first year. Because the model is composed of performance-based strategic as well as historical-based elements, it is thus far difficult to assess the implications of this reform, although the funding model will obviously have some effect on the internal distribution of resources in the universities and colleges. The institutions have already become much more concerned about the recruitment of students, dropouts and time to obtain degrees as well as research quality and number of publications.

An output-based funding model may in itself improve efficiency in teaching and research, but hardly quality (Smeby, 2003). The institutions are therefore instructed to establish their own quality assurance systems before the end of 2003. A new accreditation body, the Norwegian Agency for Quality Assurance in Education (NOKUT), has been established by the

Norwegian Storting (Parliament) in 2002 and became operative on 1 January 2003 in order to audit these systems. The consequence of not having established quality assurance systems covering a minimum standard is that the institutions are not allowed to have new educational programmes established. The Ministry of Education and Research has specified by law that there are formally three types of institutions: universities, specialised university institutions and colleges. Institutions may opt for the preferred status themselves, but have to be accredited by NOKUT to attain the status according to specific criteria. NOKUT must approve a change of status before the Ministry gives permission. However, the approval of NOKUT does not have to be accepted by the Ministry (Stensaker 2003). Academic and institutional drift processes at many colleges will therefore sooner or later probably lead to some state university colleges receiving university status.

NOKUT is also responsible for the accreditation of education programmes. All established programmes that the respective institutions have the right to offer receive accreditations automatically. Universities and special university institutions that are accredited may establish new programmes at all levels without any further procedures. State university colleges may apply for accreditation of new programmes on Master and PhD level. Private institutions are so far regulated by a separate act and have to apply for accreditation of new higher education programmes at all levels (Stensaker, 2003).

Parallel to the 'Quality Reform' the parliament decided that Norway should increase its investments in R&D substantially in order to reach the average OECD level. For the higher education system, this measure has resulted in an increased emphasis on postgraduate research training and recruitment to academic positions. Universities have also received increased research funding which they are supposed to use for strategic purposes. Furthermore, thirteen centres of excellence have been established. These centres receive additional resources from the Research Council for a period of five years.

b) natural vs. social sciences:

In the years ahead a main objective in Norway will be the strengthening of long-term fundamental research. In addition priority will be given to research in four areas: Marine research, Information and communication technology, Medicine and health care, Environmental and energy research.

c) treatment of women:

Gender equality is singled out as a perspective that should be formative while developing these four prioritised research areas. The Section for Feminist Research Policy in the Department of Research Policy in the Strategic Planning Division is responsible for providing research policy advice based on developments in gender equality in research at the national, Nordic and international levels. According to the report from the Helsinki Group for Norway, this Section for Feminist Research Policy was in charge of developing an Action Plan for Gender Equality in the R&D sector (1999-2003). The plan was adopted by the Executive Board of the Council, followed by internal and external hearings. The goals for the Action plan are:

- Strengthen the recruitment of women to fields with a low percentage of women
- Increase the percentage of women in tenured academic positions

As is stated in the Helsinki Group report for Norway, "benchmarking is a useful method to compare, analyse and improve results and processes between similar institutions in the R&D Sector in questions concerning gender equality. In Norway the method has never been used in a gender equality perspective, not until recently. The Minister of Trade and Industry announced in November (2000) a benchmarking project on the portion of women in management and board positions in the 250 largest businesses in Norway. A committee with 14 prominent women from Norwegian industry, trade, research and organisations will work out a ranking list of the 250 businesses according to their capability in appointing women to management and board positions."

SWEDEN

As was mentioned in our 1999 Report, performance-based research funding was not implemented in Sweden (p. 42) and at present time the situation remains the same. We are aware that a new funding system is being currently designed in Sweden, which is based on educational tasks negotiated between the Ministry and the individual institutions in which the three-year objectives of the HEIs are generally stated. These contracts contain the following elements (Maassen 2000):

- The number of credit points that the institutions are required to award
- The total number of FTE students (full-time equivalents)

- The fields of study in which the number of students is to increase or decrease
- The programmes in which the share of women or men is to increase
- The follow-up to be made in the Annual Report
- Special assignments.

In spite of this wider regarding of quality issues in the design of new programmes, the important elements in the Swedish contracts remain the regulation of student numbers and number of candidates and credit points awarded.

FINLAND

The 1999 Report dealt extensively with the case of Finland (pp. 42/7) and concluded by noting the intention as of 1999 to introduce a new system of research evaluation that more closely reflected the British RAE. We are pursuing our inquiries as to what happened to these plans. We are aware that in December 1999 there was launched the *Development Plan for Education and Research 2000-2004*. Within this Report, evaluation in higher education units is focused on operations, impact and quality of education and the utilisation of earlier evaluations.

The Academy of Finland, whose central role in Finnish research was emphasised in our 1999 Report, has been constantly updating and reviewing its priorities and international orientation of research activities in recent years. Its new science policy agenda (*Academy of Finland: Forward Look*) describes the changes that have been taking place and surveys the achievements and strengths of the Finnish research system and the main targets for science policy over the next few years. In the paper the Academy puts forward an eight-point agenda for developing the Finnish research system and scientific research. The main points are:

- to develop creative research environments, using mainly research programmes and the national centres of excellence programme as tools
- to expand and deepen research cooperation and specifically collaboration with sources of funding
- to fully utilise European and international opportunities in all areas of research and science policy

In June 2002, the *International Strategy for the Academy* was released, which establishes the goals for Finnish research until 2007. The new strategy is focused on a more international operation of the Academy and also wants Finnish researchers to take part in international work and promote science internationally.

There have been important reviews of government policy. For instance, the Ministry of Education has launched a series of programmes on education, knowledge and research.

The current Finnish *contractual model* was introduced in 1998. The ministry negotiates with each institution individually to determine the objectives and results that are to be achieved within a three-year period and the funds that are necessary in order to reach the objectives. The funds are allocated as a lump sum, and the institutions themselves are responsible for the distribution of resources within the university. The three-year contracts are reviewed annually in connection with the preparations for the annual budgets. Reporting and development of systems for evaluations are very important elements in the contract-based relationship between universities and the Ministry. The universities provide every year reports on the results that have been accomplished. There is also a national database (KOTA) to which the universities are obliged to report statistical data. There is no direct link for KOTA to funding, but KOTA offers basic information for the planning and budget process. The contracts are seen as the most important government instrument towards the universities, where the priorities for the coming 3-4 years are being set. It encompasses setting the targets with respect to the number of degrees within each subject field. Steering is, however, a year-round process with joint seminars and meetings between the Ministry and university leadership.

There is a direct link between the objects agreed upon in the contracts and the output-based funding system. Contrary to the Swedish and Danish systems, funding for teaching versus research is not handled by separate funding arrangements in the Finnish system. The target is that by 2003, 90 percent of the core funding of universities will be determined on the basis of targeted and actually awarded number of masters and doctoral degrees (76 per cent), the size of institution (19 per cent) and earmarked funding for graduate schools and open university instruction (5 per cent). In addition there is a an element of result-oriented funding, in that there are special funds allocated to exceptionally good research (based on an assessment from Finlandsakademiet) and exceptionally good teaching (based on an assessment from FINHEEC, the national agency responsible for evaluating quality in teaching and learning) and there are special funds for priority areas in research and teaching (Ministry of Education

2002). Examples of such projects are research in biotechnology and telecommunication and educational programmes in information and communication technology. Thus, the evaluation system in Finland is in other words partly related to funding issues.

Even though the aim with the contracts is to develop goals that may be used to assess institutional performance, apart for the number of degrees it has been difficult to develop goals which are accurate enough. Furthermore, the number of degrees is based on imperfect information since it is first of all based on targets (2/3) and not the actual number of degrees (1/3). However, no direct sanction is used even if the number of degrees falls below target, but it is taken into account when setting targets for the next three-year period. According to our correspondents, a pure formula funding system may have unintended consequences in the sense that the institutions exploit the system to increase the funding, which in this way undermines the funding system.

EASTERN EUROPE

In the time and budget available, we have not been able to collect any updating information on these accession countries. We also lack information on Germany, where we understand some changes have taken place since 1999, Austria, and Italy, where we understand no changes have taken place.

JAPAN

The administrative structure for the funding of science and technology in Japan was markedly reorganized in January 2001, when the Ministry of State for Science and Technology Policy and the Council for Science and Technology Policy (CSTP) were founded. These organizations work together to determine the national strategy for science and technology, and the policy for allocating research and development (R&D) resources. They also evaluate important national R&D projects. Included in this basic scheme is the Ministry of Education, Culture, Sports, Science and Technology (MEXT), which was formed by the merger of the former Ministry of Education, Science, Sports and Culture (MESSC) and the Science and Technology Agency (STA).

The situation of research evaluation in Japan is deeply defined by the traditional position of National Universities and their full support by the State. Poor quality and lack of international competitiveness in higher education and basic research were identified as the major causes of the recent reforms.

Kumiko Miyazaki, our contact in the Tokyo Institute of Technology, confirmed that "National universities are still part of the Ministry of Education, but are undergoing the process of becoming more independent. They will gain the status of agents in 2004 and become more autonomous." It is worth noting that 56 out of 83 national research institutes were transformed into independent administrative institutions from 1 April 2001 to increase the flexibility of administration, while also increasing autonomous responsibility.

Along with the Policies for the Structural Reform of Universities (National Universities) in 1998, MEXT prepared the *'University-based Structural Reform Plan for Revitalizing the Japanese Economy'*. These plans defined the future direction of reform in June 2001, with a view to making universities more dynamic and internationally competitive. They stipulated:

- the realignment and consolidation of national universities should be boldly pursued;
- management methods of the private sector should be introduced into national universities; and
- a competitive mechanism with third-party evaluation should be adopted by universities.

A major change in Japanese evaluation system dates from June 2001, when MEXT launched the '*Top 30 Project'*.[6] This project was designed to raise the standards of Japan's top 30 'research' universities to the world's highest levels. In each of the above-mentioned priority research areas, funds were provided to subsidize 30 graduate-level departments. The implementation process proceeds in four steps: (1) proposals by universities, (2) evaluations by panels of specialists, (3) selection of the top 30 departments in each area, and (4) provision of funding.

Under this scheme, proposals prepared by the universities are subject to a peer review by Japanese and foreign specialists who choose the top 30 departments in each priority area. The ranking scheme in this project differs from the traditional Japanese system of ranking based on the average 'hensachi' (deviation value of standardized test scores) of the applicants applying for university admission.

The long-term strategic goal of this programme is to elevate Japanese research universities to the apex of international excellence. Nevertheless, some expert academics agree on the importance of the consideration of a separate funding framework for the programme, the limitation of the government's involvement to a support role, and the need for effective internal decision-making mechanisms within universities, in order to achieve this goal. In parallel with providing prioritised investment under the Top 30 scheme, it is seen as being highly desirable for Japan to promote the establishment of networks of competence that overcome current barriers to university-industry collaboration.

The primary responsible of quality assurance of higher education in Japan is the National Government (through MEXT). Authorization and supervision by the national government is the formal base of quality assurance.

This mechanism is complemented by self-evaluation. Implementation of self-monitoring/ evaluation has been a required activity since the change of standard of 1999. According to Kumiko Miyazaki, "since the mid-1990s most Japanese universities have implemented the process of self-evaluation. External evaluation is being implemented via an organization called NIAD, the National Institute for Academic Degrees."

[6] First called the 'Tohyama Plan' and after the controversy generated by the name 'Top 30 Plan' it changed to the 'Plan of Centres of Excellence for the 21st Century of Japanese Higher Education'.

NIAD was established in 1991 and was reorganized as a new body in April 2000. In addition to its original degree-awarding functions it now works as a national organization for university research. This reform is the consequence of an earlier discussion raised by the University Council's report 'A Vision for Universities in the 21st Century and Reform Measures'. Nowadays NIAD has four major tasks:

- Evaluation of education, research and other activities of universities
- Awarding academic degrees as well as assessment and recognition/approval of programmes provided by higher educational institutions
- Conducting research on university evaluations and research on systems of academic degrees and assessment in learning adopted in other countries as well as in Japan
- Collecting, filing and disseminating information on university evaluation.

Evaluation programmes extend into three areas: university-wide thematic evaluation (UwTE), evaluation of educational activities by academic field (EEA), and evaluation of research activities by academic field (ERA).

Kumiko Miyazaki confirms that, "NIAD is currently evaluating several universities as a trial [but] real assessment will begin soon." According to her, "Japanese universities were evaluated when application to set up a new university was made, but they were not evaluated afterwards. Now, the process is about to change and Japanese universities will have to be accredited regularly (as in the US system, not like the UK system, where the outcome of the evaluation is linked to funding)."

a) use of benchmarking:

As far as we know, there is no evidence on systematic benchmarking exercises related to research evaluation or university research activity in Japan within the last 4 years.

b) natural vs. social sciences:

In March 2001, the CSTP released the *Science and Technology Basic Plan*. This report states the importance of promoting basic research and upgrading research quality and it also emphasizes the priority character of R&D. As such, it has selected priority fields, such as life sciences, information technology, environmental science and technology, and nanotechnology and materials. It is clear that preferential treatment was received in these

scientific fields of research in terms of research evaluation within the developed programmes as against arts/humanities research.

c) treatment of women

The Basic Law for a Gender-equal Society (enacted in June 1999), stipulates that the realization of a gender-equal society may be the most critical issue determining the state of Japanese society in the 21st century. This Law emphasizes the vital importance of promoting measures towards the formation of gender-equal systems in every social arena. Based on the Law, the Japanese government prepared the Basic Plan for Promoting the Formation of a Gender-equal Society (Basic Plan for Gender Equality) in December 2000. With respect to education and research activities at HEIs, the Plan suggests that such institutes incorporate gender-sensitive viewpoints, and encourage women to study a wide variety of subjects. At the same time, it suggests that advanced educational institutes promote women's participation in academic and research activities, and realize gender equality in academic circles, based on the proposals of the Working Group for Promoting Gender Equality of the Japan Association of National Universities. As far as we know, the principles established in this Plan have not yet been implemented through the adoption of a set of measures to promote the role of women in science.

TAIWAN

We have been recently informed about some intentions to introduce an equivalent of the UK RAE into the Taiwanese system. This project materialised in the 'Taiwan-UK Workshop on Research Assessment Exercise' held at the National Taiwan University at the beginning of December 2002. The National Taiwanese University and the British Council Taipei are sponsoring this project jointly. The project promotes the introduction of the UK RAE system to the research community of Taiwan, in order not only to facilitate Taiwan to initiate its own RAE system, but also to encourage future collaboration between the relevant groups in the UK and Taiwan.

While the overall standing of these intentions remains in some doubt, there is strong pressure to introduce assessment systems in particular fields, especially engineering, which accounts for one-third of NSC funding in Taiwan. On a worldwide comparison, Taiwan is 4th in ranking in its relative volume of engineering patents but only 10th in relative volume of engineering publications. Although the key focus is on the UK's RAE system as a possible model (e.g. Luo, 2002), its weakness in regard to the characteristics of engineering is recognised (unfortunately we have not been told what these weaknesses are seen as). The USA's Accreditation Board for Engineering Technology (ABET), well understood by the many Taiwanese engineers who have studied in the USA, is also a preferred model.

The National Science Council (NSC) is the major source of research project funding for Taiwan. Its programmes are relatively tightly defined through rigorous identification and review of research priorities, covering humanities and social sciences as well as natural sciences and engineering. The feasibility and originality of all projects – usually proposed by various departments of government - are evaluated by an NSC Project Committee. The Committee is also in charge of the coordination of the projects and their implementation. After project implementation, the NSC Planning and Evaluation Division is responsible for on-going assessment to ensure that each project achieves its projected goals.

The Science and Technology Information Centre (STIC) is an affiliated organisation of the NSC and performs a role akin to that of the OST in the UK. It is involved in international collaborative development of science policy and programmes and in the development of internal policy and assessment. STIC has recently worked closely with CWTS Leiden to develop a comprehensive bibliometric profile for Taiwan. The aim of the Director, Professor

Dr Hsien-Chun Meng, is to create a management information system underpinning research assessment both for analyses at institutional level and in support of international comparisons.

Academia Sinica, the prestigious national research academy, has its own research institutes and is partly akin to a national research council (although NSC is responsible for the six national research laboratories) as well as having an academy function. This involves both ex-ante and ex-post assessment and evaluation, with a lighter touch on a regular cycle. The intensive assessment process involving visiting groups, usually including international assessors, which is periodically applied to long-term investments is very much akin to the well-established system used by the UK Research Councils for their institutes.

Academia Sinica's research assessment is conducted at three levels: evaluation of research programs; development of research institutes; and performance of individual research staff. The Academia has a committee structure to evaluate research performance at each level, and makes reference to international standards by the inclusion of external referees where feasible. While the review committees are therefore composed of visiting international assessors, additional reference letters are also sought globally for each review. The standing of each Taiwan institute is regularly monitored and benchmarked through this process. This benchmark standard is then incorporated into funding and allocation decisions. For the evaluation of research programmes, the emphasis is more on the potential of the programme to contribute to the development of research capacity. Many institutes also use a process of 'retreat' when a large group of research staff will decamp with international visitors for an extensive review and exchange of both reflective internal analyses and proposals for new lines of research.

AUSTRALIA

Some reservations about the structure of research evaluation were expressed in our 1999 Report, and have evidently continued to trouble the Australian government. New strategies include its White Paper on Knowledge and Innovation (1999), the National Innovation Summit (2000), Backing Australia's Ability 2001-2005, the Higher Education Review (2002) and National Research Priorities (2002). A new implementation plan will commence in 2003.

The White Paper of 1999 (*Knowledge and Innovation: A Policy Statement on Research and Research Training*) focused on:

- Strengthening the role of the Australian Research Council (ARC) and an invigorated national competitive grants system;
- Performance-based funding for research training and research activity in universities, with allocative formulae and transitional arrangements designed to ensure that all universities are able to compete effectively under the new arrangements;
- Quality verification frameworks including Research and Research Training Management Plans, Graduate Destination Surveys (GDS) and the Australian Universities Quality Agency (AUQA, launched March 2000);
- Collaborative research programmes designed to support regional and rural communities through a Regional Assistance Package for the three-year transitional period between 2000 and 2002.

Under the new framework, universities are assisted through two performance-based block funding schemes: the Institutional Grants Scheme (IGS) and the Research Training Scheme (RTS).

The IGS supports the general fabric of institutions' research and research training activities, absorbing funding previously allocated to the Research Quantum (RQ) and the ARC Small Grants Scheme as described in our previous Report. Universities receive funding under a formula recognising their success in attracting research income from a diversity of sources (60%), attracting research students (30%) and the quality and output of their research publications, assessed through a revised publications measure (10%). The research student component of the formula is sensitive to the size and composition of the research student body in an institution, and is weighted to reflect cost differentials associated with broad fields of research.

The RTS is allocated through the Higher Education Contribution Scheme (HECS) on a performance basis, to institutions that are accredited and quality assured. The RTS provides funding according to a formula comprising three elements: numbers of research students

completing their degree (50%), research income (40%), and the revised publications measure (10%).

According to the weightings, the new scheme is more cautious with the use of research grant incomes as a performance measure (which was the base of the Research Quantum) and more sensitive to other factors related to performance (mainly student-related variables).

The performance-based funding Research Infrastructure Block Grant (RIBG) noted in our previous Report has been retained as a secondary 'block grant scheme', with a modified publication index, to fund research infrastructure. However "the scheme will remain allocated through performance-based block grants, rather than through individual research projects", to ensure that universities have the flexibility and capacity to manage their infrastructure requirements at the institutional level across all disciplines.

Nevertheless, the opinion of our contact expert, Linda Butler, is that: "recent reforms have been a backward step, concentrating more on aggregate output rather than the 'quality' of the product. While the AUQA may have been a step in the right direction, the bottom line is that a greater percentage of funding for research activities is being distributed on the basis of funding formulae. Academics' reaction to the introduction of these funding formulae has, I believe, been pretty prompt and clear."

However the changes for Australian research are still waiting for a major shock. A report from the most recent review of higher education commissioned by the Education Minister, Dr Brendan Nelson, is due out early this year. The Review is entitled '*Higher Education at the Crossroads*'. According to Linda Butler, this Report involves the arrival of new important reforms within the next few months: "commentators are predicting a significant shake-up to the sector, some even saying it will be the biggest change since the Dawkin reforms of the late 1980s. Most focus has been on the teaching side, rather than research, but research will be affected by major changes. It is expected that the funding formula[e] will be altered, though no one is predicting that they will be replaced by any alternative, though several major players in the sector have been pushing for an [UK] RAE type assessment. I'm pretty sure the publications collection as it now stands will disappear. Many of the submissions to the Review sought to have it altered or dropped altogether."

NEW ZEALAND

The system in New Zealand has long been under review, and led at the end of 2002 to the publication of the New Zealand Government's new plans (*Investing in Excellence*, Dec. 2002). The intention is to set up a Performance-Based Research Fund (PBRF) to "reveal and reward researcher excellence and excellent research". This is defined as being composed of the following:

- producing and creating leading-edge knowledge
- applying that knowledge
- disseminating that knowledge to students and the wider community
- supporting current and potential colleagues to create, apply and disseminate knowledge.

This multi-dimensioned assessment is to be gauged by a formula for funding composed of:

- quality evaluation of academics by external peer review panels: the number and FTE status of research-active staff x quality category x disciplinary cost weighting (60% weight in total funding)
- research degree completions at each subject area/academic unit: the number of research degree completions x disciplinary cost weighting x volume of research x equity loading if any (25% of funding)
- external research income earned by each degree-granting provider: the share in the total sector's external research income (15% of total)

Quality evaluation is to be assessed by 11 disciplinary-based panels, together with an overarching panel, which will classify researchers into 4 categories (very detailed indicators of the categories and eligible evidence are given in the report's Appendices, pp 53/5). The main difference from schemes such as that in the UK is that the first stage of classification will be an internal self-evaluation; subsequently the individual classifications and reviewed and if necessary altered by the panels. The first stage of internal classification is designed partly to lower the burden on the external panels but more positively to link to internal staff development procedures and to internal quality evaluation capabilities. Each individual is required to provide an evidence portfolio, which includes up to 4 research outputs but also broader outcomes such as evidence of esteem, and contributions to the development of new

researchers. The external panels will review borderline cases and conduct random sampling. Research outputs in situations where 4 outputs were not produced are to be weighted according to the Australian Department of Education, Science and Training's formula (single-authored books 5, refereed journal articles 1, patents 2, etc.). Reconsiderations of 'unfair' classifications are permitted.

There will be a 6-year gap between evaluations.

b) natural vs. social sciences:

In regard to disciplinary cost weightings, the report "was also aware that little robust information is currently available on these costs" (p 27). It adopted the distribution currently used in New Zealand for postgraduate top-ups, namely: 1 for arts, social sciences and education; 2 for science, music and fine arts; 2.5 for engineering, agriculture, medicine, etc. A problem with this is that it includes differences in teaching as well as research costs. Consideration was given to the Hong Kong weighting system and to that used in the British RAE, though these were in the end rejected, on what look like rather tenuous grounds. It is accepted that there may be some unfortunate short-term effects of introducing such a weighting system, that will have to be monitored.

c) treatment of ethnic groups and women:

As is normal in New Zealand proceedings, there is extensive consideration given to the role of ethnic variation, especially Maori and Pacific research capability. 'Cultural inclusiveness' is listed as one of the guiding principles of the PBRF. 'Maori knowledge and development' is one of the 11 review panels, and it is this which qualifies for an 'equity loading' (of 2). "Implications for women" was one of the (many) key issues identified during the consultation process (p 38), beginning with the statement: "Analyses of the UK RAE suggest that women are less likely to be put forward for assessment ...". Ways of dealing with this are rather vague but it is suggested that the self-evaluation process could allow for gaps in service etc.

d) influence of UK RAE:

At the same time, the PBRF is not intended to be a 'one-stop shop' for achieving government objectives in tertiary education, and is designed to complement other initiatives (such as strategic relevance, other funding, centres of excellence). Evidently the remit of the PBRF is wider than that of the existing RAE in the UK, though it draws on the 2001 RAE exercise as

one of its main inspirations. Its main objective appears to be to move away from a funding system which was driven primarily by student numbers rather than research quality.

The report uses the UK National Committee of Inquiry into Higher Education (1997) to outline the main objectives of higher education. In defining what is to count as 'research' both existing New Zealand definitions and the British RAE were drawn upon: "the British RAE definition provided a far greater level of detail and clarity ... and better recognised applied and industry-focused research", though the final choice was a compound of these. As noted above, there are however some significant implementation differences from the RAE.

USA

While the USA has no direct equivalent of the RAE for universities, it does conduct appraisals of research carried out by federal research agencies, under the GPRA Act of 1993. Assessments of how this was working were carried out in 1999 and 2000, focusing on the 5 largest agencies (NSF, NIH, DOD, DOE and NASA), by COSEPUP (the National Academies Committee on Science, Engineering and Public Policy). The reports reached the following conclusions:

	General policy conclusions	*Comments and reservations*
1	All agencies had tried to develop reporting procedures complying with GPRA	Some complained of extra cost and effort
2	Some agencies had used GPRA to improve operations	Some strengthened program management and communication
3	The most effective evaluation is by peer review using quality, relevance and sometimes leadership	International leadership was little used; approaches have varied
4	Oversight bodies need clearer procedures to validate and verify agency evaluations	Oversight bodies wanted better understanding of methods
5	Agencies aggregated research programs at different levels	High level aggregation made oversight assessment difficult
6	Development of human resources was often undervalued in GPRA plans and reports	Importance of educating young researchers should be explicit
7	Agencies received conflicting messages from oversight bodies about GPRA compliance	Inconsistent advice from congressional committees etc.

8	Timing requirements mean starting on next plans before performance reports are complete	A longer performance schedule could allow use of earlier results
9	Communication between agencies and oversight bodies is irregular and insufficient	Improved communication ought to reduce misunderstandings
10	Extent of use of results by oversight groups for program decisions is unclear	No indication of use of outcomes in determining budgets

A set of recommendations followed naturally from these findings, though an assessment in 2001 found that point 10 was not included among these. It is straightforward to draw some parallels between this approach to government research and an RAE view on higher education.

The COSEPUP report of 2001 noted that "This report does examine other mechanisms for analyzing research [apart from expert review], including bibliometric analysis, economic rate of return, case studies, and retrospective analyses. All methods were found to have some utility, but the people best qualified to evaluate any form of research are those with the knowledge and experience to understand its quality, relevance, and leadership, and, in the case of applied research, its application to public and agency goals." (p 12).

a) use of benchmarking:

Although leadership had been the least utilised of the three criteria, the 2001 report reviewed the application of 'international benchmarking', using the judgments of leaders in the appropriate research field. "As an experiment, COSEPUP panels [in 2000: *Experiments in International Benchmarking of US Research Fields*] performed international benchmarking in three fields – mathematics, immunology, and materials science and engineering – and found it to be faster and less expensive than procedures that rely entirely on the assembly of quantitative information, such as numbers of dollars spent, papers cited, plenary lectures delivered at international congresses, and scientists supported. / The panels also found good correlation between the qualitative judgments of experts and the results of quantitative indicators. In addition, panels concluded that quantitative measures by themselves are inadequate indicators of leadership ..." (p 15). The latter was because of difficulties in international comparisons and because the quantitative data reflected only portions of the research process (cf. COSEPUP, 2000, p 6). "In other words, numbers of papers, patents, or

citations should be used as indicators of the generation of innovative technologies, but they do not by themselves necessarily illuminate the most promising or important activities in a field." (p 16). A separate investigation of mathematics by the NSF (*Report of the Senior Assessment Panel of the International Assessment of the US Mathematical Sciences*, 1998) came to similar conclusions.

The COSEPUP study of benchmarking (2000) itemised the following "particular strengths" of international benchmarking:

- "Panels were able to identify institutional and human-resource factors crucial to maintaining leadership status in a field that is unlikely to have been identified by other methods.
- "Benchmarking allows a panel to determine the best measures for a particular field while providing corroboration through the use of different methods, as opposed to the 'one-size-fits-all' approach of some common evaluation methods.
- "Benchmarking can produce a timely but broadly accurate 'snapshot' of a field."

Specific points made included that the choice of panelists was crucial – "In particular, it is critical to include non-US participants in the selection of panelists and as panel members ..." (all panels except the oversight panel in the pilot studies did this) – and that exercises should be conducted at about 3-5 year intervals, using data for 5-10 years. Human resource development was given special emphasis. There was also some awareness that panels in particular fields might be guilty of advocacy of increased funding, which ought to be guarded against, though this might prove awkward.

An interesting technique adopted for benchmarking in all panels was the 'virtual congress' – leading experts in the US were identified in each sub-field and asked to prepare a list of speakers for a hypothetical congress to which they could invite the best in the world, from which rankings were deduced. In at least one sphere the US was clearly behind, since the method adopted for citation analysis was based explicitly on the UK OST's procedure (*The Quality of the UK Science Base*, 1997). Actually all panels identified some sub-fields where the US was not the world leader and sometimes not even among the world leaders. The mathematics panel pointed out that the US lead there was heavily dependent on recent foreign immigrants – also the materials panel concluded that US research facilities were more out of date than those elsewhere. Europe was seen as a serious future threat in immunology and

materials. It was also concluded that multidisciplinarity made benchmarking difficult in both of these latter fields.

CANADA

As was mentioned in the 1999 Report (p. 59), research evaluation in Canada has had little impact at the national level. Nevertheless since 1999 the Canadian government seems to have renewed its commitment to research. At the present time, there are three Canadian federal granting agencies – the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council of Canada (SSHRC).

The *Canadian Institutes of Health Research (CIHR)* was officially established in June 2000. It replaced the existing Medical Research Council (MRC) and has been expanded to include a virtual network of 13 institutes across the country, each dedicated to a specific area of focus, linking and supporting researchers pursuing common goals. The reorganization of the Medical Research Council (MRC) and the National Health Research Development Program (NHRDP) to create CIHR provided an opportunity to add innovative funding programmes to those traditionally available. There are two basic modes by which projects would be funded, both of which still rely on peer review for the selection of successful applicants. The first, which accounts for the majority of funding, is by *Insight Proposals*; this is also referred to as investigator-initiated or hypothesis-driven research. The other mode of funding is by *Challenge Programs*. The purpose of grants distributed in this manner is to fund research addressing certain strategic priorities identified by the Institute as requiring special attention. It is worth noting that the CIHR not only provides operating grants, but also contributes money to be used for personnel and career awards, and numerous other initiatives.

In 2000 the Government of Canada established the *Canada Research Chairs Program*. The Program was provided with $900 million to support the establishment of 2000 Canada Research Chair positions at universities across the country by 2005. The main objectives of this Program are:

- to strengthen research excellence in Canada and increase Canada's research capacity by attracting and retaining excellent researchers in Canadian universities;

- to improve, through research, the training of highly qualified personnel;
- to improve universities' capacity for generating and applying new knowledge; and
- to ensure the effective use of research resources through strategic planning by the institutions as well as the inter-institutional and inter-sectoral collaboration, as appropriate.

The allocation of Research Chairs to a university is based on how much federal grant agency funding a university has received, including funds received by any affiliated research institutes and hospitals. There are two types: seven-year renewable Tier 1 Chairs and five-year Tier 2 Chairs, renewable once. There will be strong linkages between the Canada Research Chairs Program and the programmes of the federal granting agencies. In fact, in their nominations, universities have the opportunity to include a request for infrastructure support from the Canada Foundation for Innovation (CFI).

REFERENCES

Geuna, A., Martin, B. et al. (1999) 'Resource Allocation and Research Performance: the Assessment of Research'. A study carried out for HEFCE, SPRU, University of Sussex.

Ireland

ICSTI (1999) 'Benchmarking School Science, Technology and Mathematics Education in Ireland Against International Good Practice'

ICSTI (2001) 'Benchmarking Mechanisms and Strategies to Attract Researchers to Ireland'

ICSTI (2002) 'Measuring and Evaluating Research: report of the ICST task force', ICSTI/Forfas.

IRCSET (2002) 'Review of the framework for national policy in research and technological development', Embark Initiative/IRCSET.

Helsinki Group on women and science (2001) Report on women and science: Review of the situation in Ireland; Irish report by Jacqueline Allan, FORFAS (July).

Netherlands

VSNU (1999) 'Quality Assessment Made to Measure'; Protocol for the External Assessment of Educational Programmes (2002-2005).

VSNU (2003) 'Standard Evaluation Protocol (2003-2009) for Public Research Organisations'.

Helsinki Group on women and science (2002) Report on Women and Science Review of the situation in Netherlands; Dutch national report by Ilja Mottier, Ministry of Education, Culture and Science (May).

Flanders

Debackere, K. and W. Glänzel (2002), 'Using Bibliometric Approach to Support Research Policy Making: The Case of the Flemish BOF-key', mimeo, K.U. Leuven, Steunpunt O&O Statistieken.

Helsinki Group on women and science (2002) Report on Women and Science Review of the situation in Flanders *(Belgium)*; Flemish national report by Koen Verlaeckt and Lut Bollen, Ministry of the Flemish Community (January).

Switzerland

SSTC Document 4 (2002) 'Structural Reform of the Swiss Higher Education System'

SNSF (2003) Guide 2003 'National Centres of Competence in Research'.

Denmark

Ståhle, B. (1999): Alder, køn og rekruttering i dansk universitetsforskning .

Langberg, K. (2002) 'Benchmarking and ranking of publications. CASE: Economics, Management and Finance', Working papers 2002/9: http://www.afsk.au.dk/ftp/workingpapers/WP2002_9.pdf

Smeby, J.-C. et al. (2003) 'Contract Arrangements in the Nordic Countries - solving the efficiency – effectiveness dilemma?'; submitted to *Higher Education in Europe.*

Helsinki Group on women and science (December 2000): Report on Women and Science Review of the situation in Denmark; Danish national report by Merete Reuss and Karin Kjær Madsen, Ministry of Science, Technology and Innovation.

Norway

Smeby, J.-C. et al. (2003) 'Contract Arrangements in the Nordic Countries - solving the efficiency-effectiveness dilemma?', submitted to *Higher Education in Europe.*

Stensaker, B. et al. (2003) 'Tertiary Education and Management', Volume 9.

Helsinki Group on women and science (2001) Report on Women and Science Review of the situation in Denmark; Norwegian national report by Elisabeth Gulbrandsen, Research Council of Norway (December).

Sweden

Maassen, P. (2000) Models of Financing Higher Education in Europe. In *Report from the Nordic Theme Conference – April 2000 Reykjavík, Island – Financing Higher Education*. Tema Nord (2000:539), Copenhagen: Nordic Council of Ministers.

Japan

CSTP (2001) 'Science and Technology Basic Plan'

MEXT (2002) 'Academic Frontiers Student Exchange Promotion Program in 2003-2004'

MEXT (2000) 'University-based Structural Reform Plan for Revitalizing the Japanese Economy'

Taiwan

Luo, R.C. (2002) 'From Engineering Accreditation to Taiwan's International Links on Research Assessment', (President), National Chung Cheng University.

New Zealand

Ministry of Education (2002) 'Investing in Excellence: the report of the Performance-Based Research Fund working group'

USA

COSEPUP (2000) 'Experiments in International Benchmarking of US Research Fields'

COSEPUP (2001) 'Implementing the Government Performance and Results Act for Research: a status report'.

NSF (1998) 'Report of the Senior Assessment Panel of the International Assessment of the US Mathematical Sciences'

LIST OF CONTACTS

Ireland

Dr. Rhona Dempsey
Manager, STI Indicators,
Science, Technology and Innovation Division,
FORFAS, Ireland

Dr. Aidan Kane
Director Centre for Innovation & Structural Change (CISC)
Department of Economics
National University of Ireland, Galway
Ireland

Netherlands

Prof. Don F. Westerheijden
Center for Higher Education Policy Studies (CHEPS)
University of Twente, Netherlands

Dr. David F.J. Campbell
Institute for Interdisciplinary Studies (IFF) at Austrian Universities, Vienna

Guy Neave,
Professor and Scientific Director,
Center for Higher Education Policy Studies,
Universiteit Twente,
Netherlands.

Belgium

Prof. Koenraad Debackere
Leuven Research & Development,
K.U. Leuven,
Leuven, Belgium

France

Bastiaan de Laat
Director of Technopolis
France

Switzerland

Dr. Stefan Bachmann
Swiss Priority Programmes /
National Centres of Competence in Research
Swiss National Science Foundation, Bern

Denmark

Finn Hansson
Lektor
Institut for Ledelse, Politik og Filosofi,
Denmark

Norway

Dr. Svein Kyvik
Norwegian Institute for Studies in research and Higher Education
Oslo, Norway

Dr. Jens-Christian Smeby
Norwegian Institute for Studies in Research and Higher Education (NIFU)
Hegdehaugsvn. 31, N-0352 Oslo, Norway

Japan

Professor Yasunori Baba
The University of Tokyo
Research Center for Advanced Economic Engineering, Japan

Prof. Kumiko Miyazaki
Tokyo Institute of Technology
Japan

Taiwan

Jonathan Adams,
Director Evidence Ltd,
Taiwan

Dr Yen-ling Lee,
Department of Asia-Pacific Industrial and Business Management,
National University of Kaohsiung
Kaohsiung 811, Taiwan

Australia

Prof. Linda Butler
Research Evaluation and Policy Project
Research School of Social Sciences,
Australian National University
Australia

New Zealand

Peter McKinlay
McKinlay-Douglas and Co.,
Tauranga, New Zealand

1.2 Changes in national assessment policy. 1.3 External influence over national educational policy. 2 Framework.Â Emerging research suggests that international standardised assessments are having an impact on national assessment policy and practice. PISA is being integrated into national policies and practices on assessment, evaluation, curriculum standards and performance targets; its assessment frameworks and instruments are being used as best-practice models for improving national assessments; many countries have explicitly incorporated and emphasise PISA-like competencies in revised national standards and curricula; others use PISA data to complement national data and validate national results against Assessment Practices: Studentâ€™s and Teachersâ€™ Perceptions of Classroom Assessment. CHAPTER 1. University of Massachusetts Amherst.Â In the Western countries at present, students are encouraged to fully participate in classroom activities.Â Similarly, other researchers agree that the core features that characterize formative assessment are that it impacts the quality of teaching and learning, and it engages students in self-directed learning environment (Chappuis & Stiggins, 2004).Â In addition, Harlen (2006) justifies changes in assessment practices, to be used in four purposes: diagnostic, formative, summative, and evaluative. The transformation of assessment practices, according to Herrera et al. The IBE programme of co-operation in â€˜research and studiesâ€™, â€˜training and capacity buildingâ€™, and â€˜the ex-change of information and expertiseâ€™ is based upon two major assumptions: l Although different Member States of UNESCO have very uneven and heterogeneous experiences in the de-sign and adaptation of their educational content, there is room for beneficial exchanges between countriesÂ He shows how current practices in science teaching are based on a number of myths or assumptions that are not always justi-fied. He examines the existing obstacles to change and ends by suggesting some elements of remediation â€˜for be-yond 2000â€™.Â Gregorio summarizes some of the main challenges that countries in Asia and the Pacific have to face in science education.