EUROPEAN COMMISSION

Gender and Excellence in the Making

2004 Directorate-General for Research EUR 21222

Part 1 - Synthesis report on the workshop

Margo Brouns and Elisabetta Addis

Synthesis report on the workshop*

Introduction

The issue of gender and excellence represents a relatively young research field in which questions, methods and answers are developing rapidly. Scholars from very different fields and using very different methodologies gathered together for the workshop. Some participants were experts in measuring scientific output on a large scale, while others were primarily oriented towards the multifaceted issue of gender and science. Qualitative case studies on micro-politics, text analysis on placement reports, experiments with men and women judging each other, and extensive surveys on success scores of men and women regarding research applications were examined together in an effort to unravel the different aspects of the complicated issue of gender and scientific excellence. In this context, exchange of ideas and research results were expected to be very fruitful: sometimes dissension is a major source of inspiration for scientific developments.

In this synthesis report, we have tried to summarise the most important factual elements and discussions from the workshop, and streamline the conclusions and recommendations arrived at in the final session. This synthesis report opens with a description of the problem, the dilemmas regarding existing measurements of scientific excellence, followed by a chapter on masculinity and male bonus, and the social dynamics regarding development and measurement of scientific excellence. The next chapter investigates mechanisms causing indirect gender bias and, finally, the report will draw some general conclusions and make some recommendations.

Framing the problem

> *"The low female presence at the highest levels of the scientific hierarchy is an indicator of the inability of research institutions to follow changes in society, such as the increase in women in higher education, which in turn highlights the dysfunction of a system for the evaluation of scientific excellence that has not abolished or weakened the old boy network of co-optation."* (Palomba, this volume, p. 5)

The last 15 years have witnessed unprecedented interest in the presence of women in the sciences. The career paths of men and women in almost all branches of scientific endeavour have been investigated (e.g. Sonnert and Holton 1996a), the aim being to unravel the complex interactions between institutional arrangements and personal preferences that might explain the under-representation of women, especially at the highest levels. This research has produced new knowledge and understanding of the 'gender of careers', as well as divergence in interpretations of the current situation and the appropriate strategies to deal with any problems. One issue that remains central to current research is whether achievements of men and women are assessed on the same basis, and it is this question that underpins the contributions to this report.

In 1997, Wennerås and Wold published their ground-breaking *Nature* article on sexism and nepotism in the peer review of research grant applications to Sweden's Medical Research Council (MRC). The article showed that the peer review system is not as 'neutral' as it claims to be. Male applicants and researchers with an affiliation with one of the evaluators were more successful in their applications to the MRC for postdoctoral research grants. The article concluded that whilst the quality of the proposal was an important factor in assessing the scientific competence of research grant applicants, so was the gender of the applicant, as well as his or her affiliation to one of the members of the evaluation committee. These applicant characteristics proved significant in subsequent assessments of research grant applications in Sweden[1].

This evidence of gender bias was particularly disturbing because it contrasts with one of the scientific community's core beliefs about its own internal governance. Decision-making should be based on meritocracy, hierarchy on

(*) This synthesis report reflects primarily the view of its authors. Participants have had the opportunity to review the first draft. Many of them have sent their comments, for which we are very grateful. The authors are, nevertheless, responsible for the final synthesis report.

individual performance in furthering scientific inquiry. This belief is rooted in the heart of the scientific ethos, connected with the struggle of science to liberate itself from theology and other societal powers (Merton, 1942). Originally, it was a plea for autonomy of the scientific community and democratic self-government within the scientific community. Only the accomplishments of the individual scientist should count, and only the scientific community is capable of evaluating accomplishments. A scientific enterprise cannot afford to base the acceptance or rejection of scientific claims on the personal or social attributes of their protagonists (Gupta, Kemelgor, Fuchs & Etzkowitz, this volume). Wennerås and Wold's (1997) results questioned this core belief and the validity of scientific selection procedures. A science that is oriented by non-scientific judgments on the performance of some scientists lacks the fundamental quality of objectivity. In this context, the supposition of attributing 'excellence' mainly and mostly to male scientists becomes problematic for all scientists.

Methodology and evidence of gender bias

The Wennerås and Wold publication prompted other investigations into research funding. However, the results from studies in Great Britain (Wellcome Trust, 1997) and Germany (Boehringer Ingelheim Stiftung, 1999) found the success rates of men and women to be rather similar. The authors conclude, therefore, that there was no gender bias in the systems investigated. However, they did not research the files of the applicants in order to validate the conclusions. A Dutch study showed that even if, on a general level, success rates of men and women are similar, gender bias can still be active. Due to processes of selection and self–selection, it can be expected that on average women applicants are better qualified than men.

Those who question the existence of gender bias often point to other circumstances in which gender bias does not occur as confirmation of meritocracy as one of the central values of the scientific system. This caused a lively debate on the evidence of gender bias during the workshop, and it did not end in consensus. As John Stuart Mill already stated in the 19th century: "*In every respect the burthen is hard on those who attack an almost universal opinion*".

General issues

However, there are more reasons to question current definitions and measurements of scientific excellence. The peer review system has long been the subject of debate within the scientific community. For decades, most scientists published their views without formal peer review and peers published their criticisms openly. However, at the turn of the 20th century, the breadth of scientific endeavour widened to the extent that editors of journals and fellows of societies began relying increasingly on the advice of specialists. The editorial system used by leading journals such as *Nature* runs on similar lines to those used more prevalently until about 20 years ago, when many journals relied on editors to make decisions without the formal refereeing of papers. Henneberg (1997) argues that peer review, as now practised by most scientific journals, might stifle scientific communication and slow down the advancement of knowledge.

According to Cozzens' input into the discussion, the long-lasting debate on 'objectivity' and 'fairness' of assessment systems ended more or less in the statement that although the system is not perfect, it is the best available (Cole and Cole, 1985). However, with the growing interest in gender issues in the field of research policy, the picture seems to be changing: the system is not only imperfect, it may even be hindering women in establishing scientific careers. Merit and talent are not sufficient conditions to become a successful scientist. Resources, time, social networks, encouragement – unevenly distributed between the sexes – are necessary prerequisites.

Gender issues

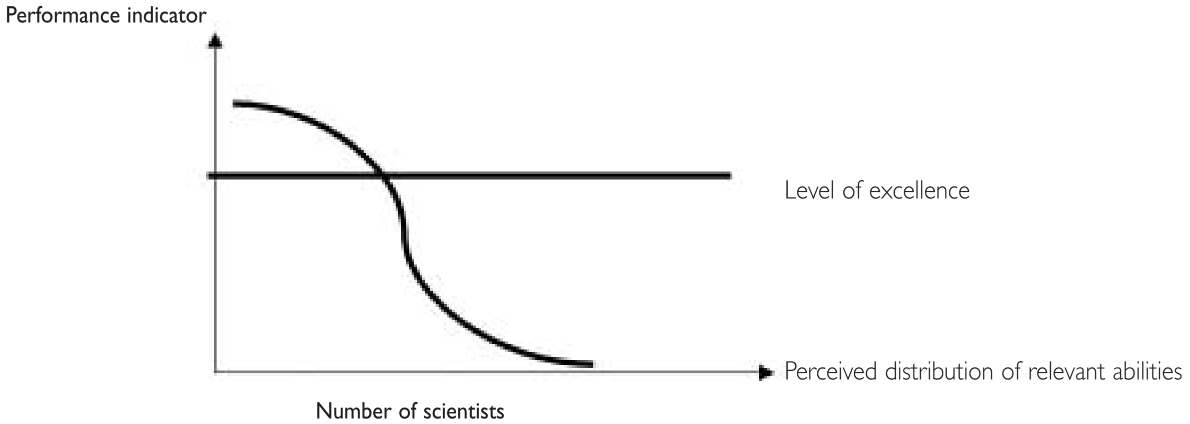
Although the workshop focused on 'excellence' – i.e. the very high levels of scientific performance – the mechanisms that appear to prevent women scientists from achieving excellence partially overlap with those that prevent women from rising from any step of the academic ladder.

In some national systems, direct gender discrimination has been prohibited for many years, but this does not mean that it is not practised. Research has found gender to be a strong and meaningful category in all aspects of social life, and thus also in the sciences.

> **Gender bias**
> Gender bias is the often unintentional and implicit differentiation between men and women situating one gender in a hierarchical position to the other, as a result of stereotypical images of masculinity and femininity steering the assessment and selection process or the gendered structure of the scientific system. Explicit gender bias is prohibited, but still exists – discriminatory practices considering recourses seemed partly to explain the under-representation of female scholars at the highest positions (MIT 1999).

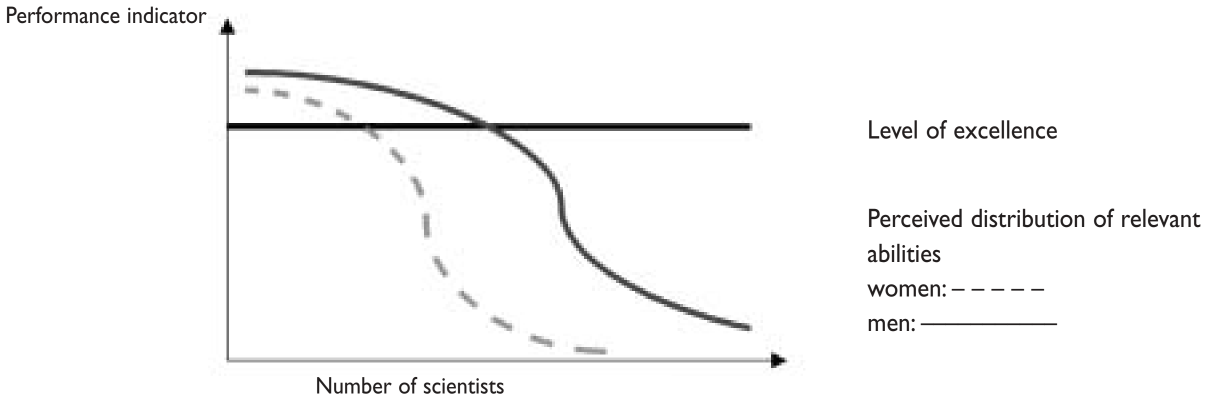
Participants in the workshop presented evidence from their research both of direct and indirect gender bias. Direct gender bias refers to the ways scientific competences are attributed to men and women, see for instance the report by Foschi (this volume) which clearly points to the existence of a double standard in the assessment of task performance by men and women. Indirect gender bias refers to the unintended negative effects the organisation of sciences and scientific assessment has on the opportunities and challenges for women scientists. An important issue is that of the hegemonic position of the 'hard sciences' vis-à-vis the humanities and social sciences, in combination with the relative absence of women in the hard sciences.

Before a more extensive discussion of the multifaceted gender bias, we should first turn to more general problems related to measuring scientific excellence. Simplism is surely one of them. The dangers that derive from using an over-simplistic model of excellence are illustrated by the following set of graphs. (*)

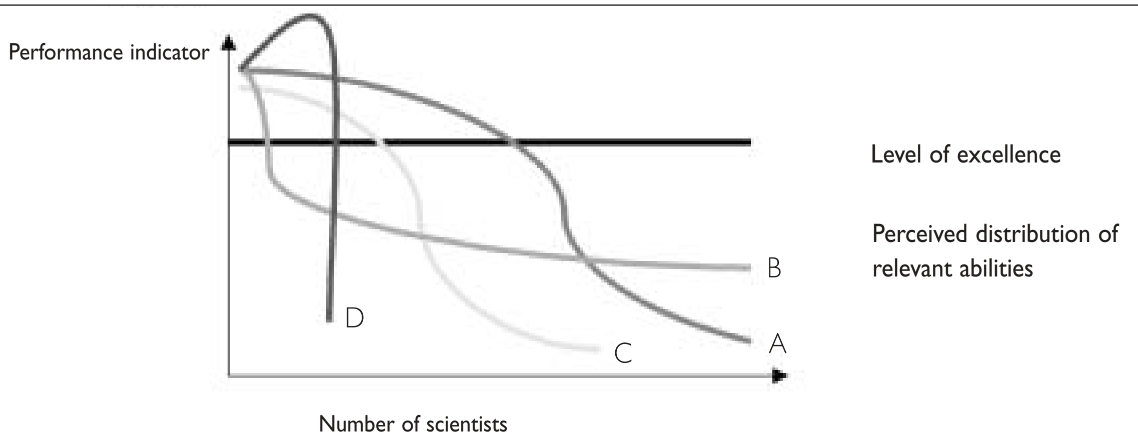


Some may think of excellence as if it was an easily and objectively measurable variable. There is a performance indicator, related to publications and other accomplishment, and there are scientists. Those who reach a certain level of the indicator, as established by the scientific community, are "excellent". More male than female scientists appear to attain that level.
Notice however that….

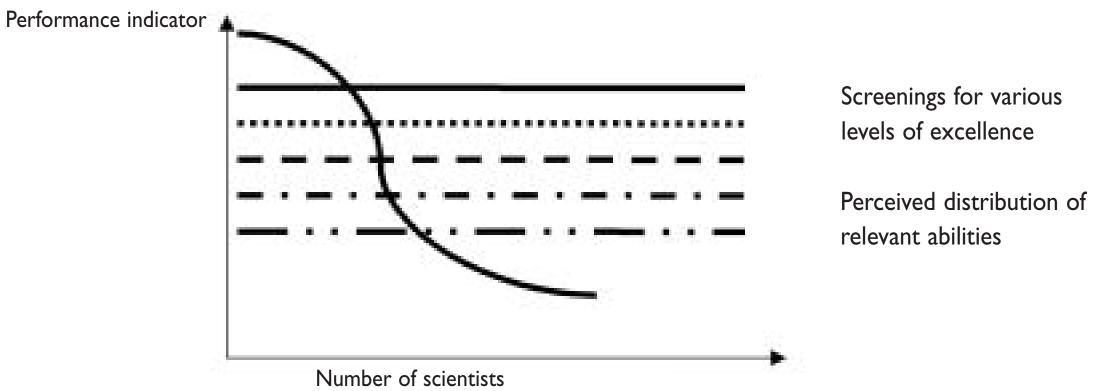
(*) Graphs are Elisabetta Addis's closing contribution to the workshop. Names in parenthesis are not direct quotations. They are the contributors to this volume whose presentation inspired the graphs. They are not responsible for the interpretation. Only Elisabetta Addis is, and she thanks them all.



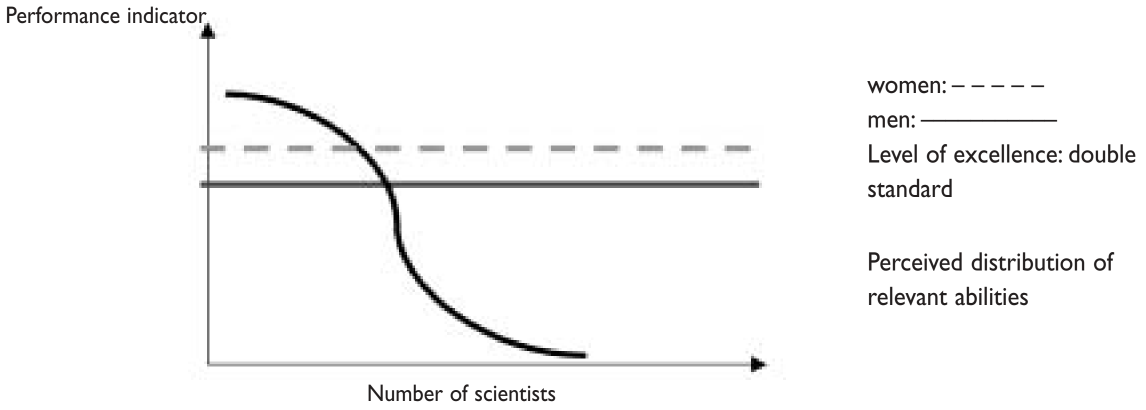
…for purely statistical reasons, even if abilities relevant to excellence have the same distribution among M and F in the population at large, **a low F/M ratio in the population of scientists** – related to the triple burden, (Kemelgor, Gupta, Fuchs, Etkowitz) – **produces a low F/M ratio among the "excellent"**. This may lead to the wrong conclusion that women are less apt than men to be "excellent".



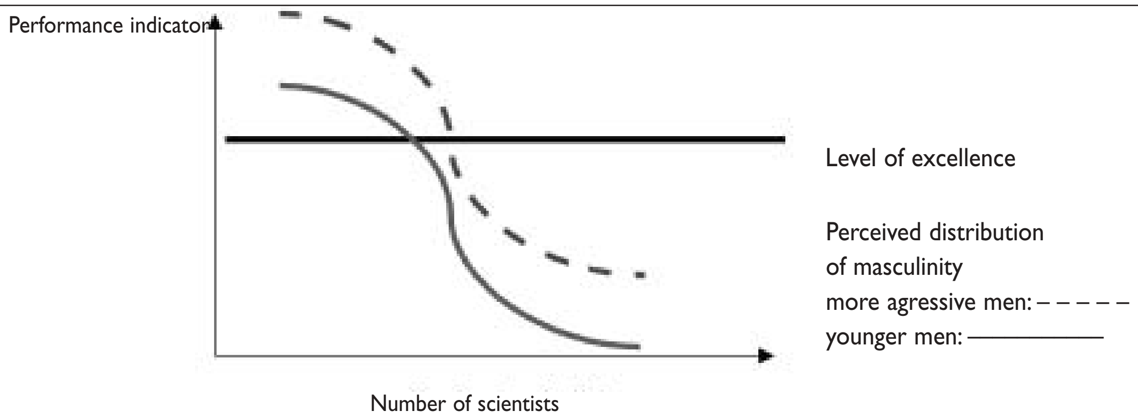
… Notice also that the distribution of abilities is not a natural fact. It is a social construction (Griffin). Its shape varies depending a) on which abilities are counted as relevant b) on what is the indicator of performance chosen c) on who is counted as a member of the scientific community. Therefore, the distribution of excellence to some appears as the line A, to some as the B line, and to some as the C or D line, depending on what abilities they factor in…



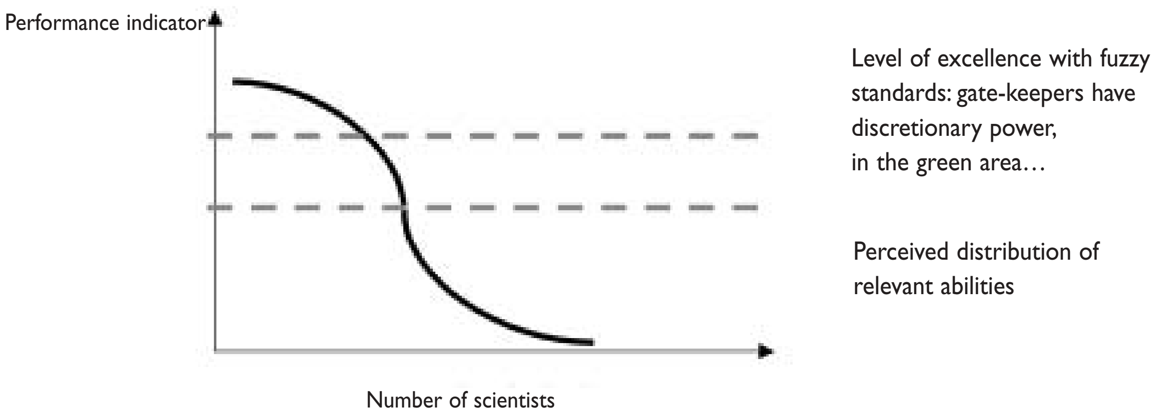
Moreover, the top level of excellence is attained after passing through a number of screenings. If after any level the F/M ratio is lower than unity, this will produce a cumulative effect that will make women scarcer and scarcer on the way to the top. (Palomba). For each of those screenings, there are "gate-keepers"(Husu), and their qualifications may be questionable. They may look for some characteristics that are unrelated to the pursue of knowledge (e.g. unconditional availability of time, assertiveness…) or, even unawarely, marginalize and devalue women's contribution (Thorvaldsdóttir).



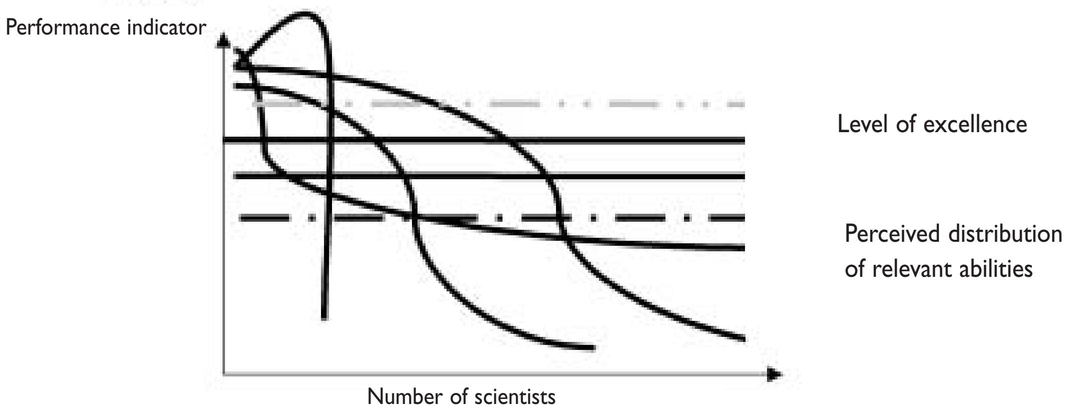
Gate-keepers may assess excellence using double standards (Foschi); Standards may be set higher for women, for scientists who come from disadvantaged backgrounds (Cozzens), or from relatively non-hegemonic nations (Blagojevic).



…and sometimes gate-keepers are just looking for a ranking according to a set of qualities we can better define as "masculinity" (Hearns). Some men, many women and often younger people are less endowed with it….



If the standards to be achieved in order to be judged excellent are fuzzy, rather than been clearly stated, it may favour the application of double standards. For some people it is enough to pass the lower dashed line, but some others need to overcome the higher dashed line. The same may happen when gate-keepers are not accountable and procedures are not transparent.



If we consider all that, the model of the first box deconstructs in a rather more complex scheme. Excellence as we see it today is just one of many possible consensuses about what excellence is. This consensus is shaped by gender relations in the scientific community and in society at large. But the standards used may be different, and the distributions may be perceived differently. Excellence can change. Thus excellence becomes a contested terrain. The effort to measure excellence is also a battleground. Existing measures, like bibliometrics, are not gender-biased, but this is not the same as saying that there is not structural gender-bias in the larger environment (Feller). The effort to establish criteria and build indicators that take into account the difference in men's and women's lives and abilities is an effort to redefine excellence so that excellent people of both genders may contribute to science.

Definition and measurement of excellence: the mainstream

Scientific excellence is essentially difficult to grasp. In his seminal work on the ethos of science, Robert K. Merton (1942) stated that for science to be fertile and productive, scientists must be judged only by their work, and win status and membership within the scientific community on that criterion alone. From this perspective, the scientific forum is the best institution capable of evaluating the results of research. Peers should, therefore, assess the quality of research proposals and products. Because it is based on these principles, the peer review system professionally produces 'certified knowledge'. Disinterestedness and the ability to be objective are cornerstones of the scientific ethos.

Bibliometrics

The most important device used by peers when reviewing research results, which most academics are willing to recognise as 'unbiased', is the number of published articles and books and their influence. However, there is a tension between reliability and validity. Indicators that are measured easily and unequivocally – and provide a reliable way of counting – are not necessarily the most valid. In other words, bibliometrics are not necessarily the best indicator of scientific quality, as argued by Feller in this volume. What can be measured is not the full potential of quality but only representations of quality. Bibliometrics are used as a proxy for excellence, quality, and ability. The quantitative is a reduction of the qualitative, which is not easily measured in an objective manner. Moreover, the connection between short-term publication and long-term scientific impact is rather weak. Early measures may not be an accurate predictor of the long-term impact of a scientific discovery: Wittgenstein would not have survived such a system (Dummet 1991 cited in Feller in this volume).

The validity of the Science Citation Index regarding scientific excellence is limited as it rarely includes sources in languages other than English, and covers only a minority of the scientific journals in humanities and the social sciences (see chapter below). The system of judgment employed in bibliometrics privileges well-established fields with long-standing publication traditions and clear boundaries. In these research fields, scholars have ample opportunity to expand their scientific activities along existing research lines, to participate and exchange within an established scientific community, to get funding for their research and to publish in several journals, ranked in prestige from high to low. For new research fields, these facilities – resources, publication sites – still need to be developed, and measuring scientific quality causes some difficulty – especially if quality is measured using bibliometrics as these fields are not yet visible enough. The same holds true for scholars active in relatively narrow fields – for example, certain languages – who are easily underestimated when it comes to measuring quality by citations and publications, which certainly applies to scholars engaged in gender issues (Griffin, Rees and Mottier, this volume). Bibliometrics privileges clear and intelligible sets of rules that successfully translate scientific publishing into economic rewards, which is not the case in issues and countries at the periphery, as Blagojevic argued.

Another major and still increasing problem caused by mainstream measurement of scientific quality is the way individual scientists react by producing more and more publications. An unintentional performative effect of bibliometrics as the standard measure for scientific excellence may be that there are now even too many publications in certain areas. The way scientific excellence is measured creates a specific atmosphere in which competition leads to high numbers of publications but not necessarily to good science. "*Publication numbers themselves can be an outcome of a certain form of masculinity*", as Hearn stated during the workshop. The competitive atmosphere tends to stimulate the individual performances of scientists competing with one another. The result is a stream of publications.

An additional problem with bibliometrics is the result of the same computer technology that made bibliometrics possible. Writing an article in 2003 was different from writing it in 1993. The availability of computers with more capability, better word-processing programs, improved printing facilities, and access to multiple bibliographic sources through the web has led to the possibility of writing more articles using the same amount of effort. As a result, scientific production has increased massively. There has been a veritable inflation of literature, while the information-processing capabilities of humans has stayed the same.

Since the 1970s, a lot of research has been done on the issue of gender and publications. On average, women tend to publish fewer articles than men (see the text box). This so-called *productivity puzzle* is a fascinating issue. Recent publications clearly show that productivity appears to be related to academic rank. The lower productivity of women can be explained by the fact that they are working at lower professional ranks than men. Within the same category, it seems that there is no significant difference by gender (Bordons et al., 2003). In addition, there are important differences between the scientific fields in terms of women's participation and of publication rates and citations. Discipline-specific publication traditions can explain the existing gender differences in productivity.

Women and publishing

According to Schiebinger and Valian, there is some evidence that women tend to publish fewer papers, with each paper being more substantive. On average, papers published by female scholars are cited more frequently than papers by more 'productive' male scientists (Schiebinger 1999, Sonnert & Holton 1996b, Long 1992, Zuckerman 1987, Nilsson 1997; Feller this volume). Palomba (this volume) reflects on this issue: differences are not so much caused by numbers of articles but by other aspects, such as type of publication, language or type of specialisation (p.3). However, Palomba found no difference in impact factor: her data show that publications by women are as influential as those by men.

An alternative explanation for gender differences in publication rates emphasises family responsibilities. During the workshop, Palomba presented research which showed that there is a family effect on productivity: the publication peak for men is earlier in their careers than for women. A lack of reliable data comparing men and women working under similar conditions (age, experience, etc.) makes it difficult to draw conclusions on gender differences in publication rates across disciplines and age. Nevertheless, these results were not confirmed by research carried out by Sax et al. (2002) which shows that factors affecting faculty research productivity are nearly identical for men and women, and family-related variables, such as having dependent children, exhibit little or no effect on research productivity.

During the meeting, Sandström suggested that the use of bibliometrics benefits women scientists. According to his data, women are the most productive (number) and influential (citation index) researchers. From this point of view, women could profit from the more 'objective' measurements of scientific excellence by counting publications and citations. Bibliometrics are reliable in the sense that counting numbers of publications is not influenced by who is counting, thus avoiding implicit subjective judgments at this level.

However, bibliometrics may still be gender biased if the measurement and the underlying criteria reflect the scientific activities of men and women differently – for instance, a citation index that is primarily focused on the natural sciences and covers only 20% of the journals in social sciences and humanities has limited validity in measuring the achievements of female scientists. In these cases, although the counting may in itself be exact, the underlying processes that lead to citation may reflect practices so different as to render the counting meaningless or biased.

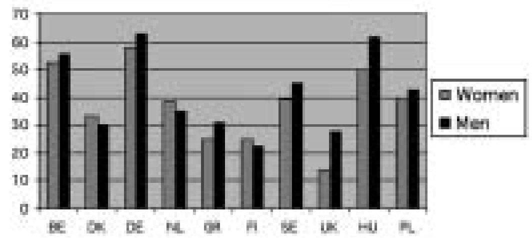
Peer review

There was general consensus among the participants in the workshop that excellence is not an 'universal fact' or a 'natural given', or a 'supra-disciplinary' fact. It is a social construction and, as such, it is open to many kinds of bias. As suggested by Addis in her contribution, it would be wrong and misleading to treat excellence as if it was a simple, easy-to-measure quality, like height or speed. Instead, it is a composite of many skills – carefulness, originality, clarity, complexity, etc. – and is achieved through a process of training, networking, accumulation, and resources. The judgment of excellence depends on the importance attributed to each of these characteristics. Therefore, it is dependent on the people judging and the standards they set for excellence.

The most obvious difficulty is the evaluation of original, innovative research. Innovation is not always recognised immediately, and may sometimes even be rejected as 'bad science'. Other problems arise because of the idiosyncratic character of the judgments. According to Feller (this volume), there seems to be a lack of attention to or endeavour in the construction of alternative measures. The challenge for future work assessing scientific performance and excellence is to develop metrics that better capture the dynamics of scientific discovery, as well as encompass the array of societal objectives that led to the initial public policy decision to fund the research.

Research funding success rates in EU Member States 2001

In some countries, women are more likely to submit successful research funding applications (Denmark, Finland, the Netherlands), but in others they are less successful than men (UK, Germany, Sweden). The diversity of grant application procedures between countries is strong, even stronger than the diversity between the sexes.



(Source: She Figures 2003: 75)

However, similar success scores of men and women do not provide a final answer to the gender bias issue. Due to the self-selection processes, we might expect female applicants to be on average better qualified than men. Unless they are really qualified, women do not bother to apply for a subsidy (Brouns, this volume).

In addition, the presupposed disinterestedness of the peers and the objectivity of the system are the subject of debate. A standard called 'objective' can be interpreted differently: one, because similar results are achieved when applied to similar situations, and two, because there is consensus that it really is capable of doing just that. However, as Griffin noted in her paper, there is no golden standard – all those participating in the decision-making process must agree on a standard and/or rules to decide whose standard will prevail in order for a decision to be made. Thus, choosing the standards, proper indicators, and devices capable of measuring the standards is a point of contest between different viewpoints. The 'objectivity' of the final decision is the result of negotiation.

'Similarity' seems to be a major aspect steering the evaluation process. Although existing research is ambiguous, there is some evidence that peer reviewers prefer proposals that are similar to their own work (Guetzkow, Lamont & Mallard [in press] cited, with permission, in Griffin in this volume). Knorr-Cetina (1999) has called practices based on similarity 'epistemic cultures', a primary orientation and research styles characterising research groups and research fields. These preferences function as the frame of reference for judging 'excellence'. This implicit cloning mechanism(2) limits the chances of research proposals and publications that do not fit in with the traditions. Moreover, evaluators tend to overestimate the accomplishments of scientists with an established reputation, whereas unknown researchers meet more reserve. This so-called Matthew Effect (Merton, 1968) – achievements are attributed to the more famous researcher – has been documented extensively. The gendered variant, called the Matilda Effect, has also been documented: achievements of female researchers are frequently attributed to their male colleagues or otherwise minimised and underestimated (Rossiter, 1993; Stamhuis, 1995).

Some participants in the workshop argued that the criticism refers to a conservative approach towards the sciences. The scientific profession has changed dramatically and has become internally segmented. Career tracks are fragmented because scientists and scholars are active partly inside universities and partly inside national labs. Scientific careers are characterised by a variety of activities but only a few become visible; many of them stay unnoticed by the assessment systems. It is impossible to address all the dimensions of an individual's professionalism. In this case, research is about 'good' and not about 'best' (Laredo, this volume).

Masculinity, male bonus and cultural interpretations of gender

Some papers indicate that the academic career system is based on the traditional male model of labour market participation. A scientific career presupposes long working hours, creating a rather one-sided work-life balance for many researchers – especially those with family responsibilities – that both men and women find hard to live up to. The ideal type is essentially a male model of practice, full-time devotion, emphasis on early achievements, and exclusive identification with science, without any other social obligations. In this context, equalising resources is not enough to create a situation of equity, as the transmission of resources into a career is not the same for all people (Cozzens, Rees in this volume). This situation has been described by the catchphrase 'male bonus' (Thorvaldsdóttir): the problem is not so much that women encounter discrimination as such, but that people – men and women – who resemble those who are in powerful positions and behave according to masculine traditions of full-time devotion and competition enjoy a bonus that allows them to be assessed as better scientists. The winner seems to take all.

Many institutional evaluation guidelines acknowledge the impact of a part-time job on productivity, for instance, the recently implemented British Research Assessment Exercise (RAE), but at the symbolic level the idea of 'the scientist' described above is still active. Despite the guidelines, part-timers and 'women returners' (those returning to academic work after a career break) are vulnerable to a negative judgment due to limited publication records. Although, in the UK, 50% of the part-time workers are male scientists, it is generally the female part-time workers who experience a negative impact on their positions; they are 'time poor' (Rees, this volume). The male part-timers are generally approaching retirement and thus do not fear for their careers. In fact, they are 'time rich'.

Men are over-represented in the senior positions of academia, especially in Europe. This male dominance tends to lead to women having increased visibility as women. As Hearn stated during the workshop, only women are seen as having a gender. "*Most analysis and policy development in research and academia, and often even that which is concerned with gender, continues not to gender men explicitly and not to make explicit men's part in the problem of gender inequalities*."

(Hearn, this volume, p.1) Mechanisms of discrimination against women are discussed frequently, but what about the way men are treated? There is still much research to be done on this issue, both in academia and more generally. There is a complex interconnection between men's practices, academic practices, and managerial practices. This creates a highly gendered atmosphere. As a result, most applicants and evaluators are men, and they also serve as gate-keepers for new positions. Following Essed & Golberg (2002), Hearn argues that the recruitment of new managers is closely related to processes of cultural cloning, pointing at an often unintended preference by men for men or, as Kanter stated some decades ago, homosocial reproduction (Kanter, 1977). The similar-to-me effects implicitly influence assessment and selection procedures. According to some of the participants in the workshop, this privileges the contribution of men and diminishes that of women. This might also produce competition between men and make them rivals as well as supporters.

Women scientists seem to encounter trouble in becoming part of loose networks, subtly excluded even by colleagues who are not explicitly sexist in any way. One possible explanation of this fact is that men competing against each other can expect large honour gains when they win and only small ones if they lose. In competition with a woman, the picture changes: men do not want to compete with women because the gains from winning the competition are relatively small, and the risk related to losing the competition is high, because this implies large honour losses (Addis, this volume). As a result, men treat women differently from the way they treat men, and women remain 'the others'. Under these circumstances, it is far easier for men to gain scientific credibility from an overwhelmingly male scientific forum than it is for women.

Based on an analysis and close reading of texts considering university placement committees, Thorvaldsdóttir's paper describes several mechanisms of gender bias in selection procedures. Gender-biased language may be used, as well as gender-biased characteristics, such as deprecation and silencing, and negative criticism, fault-finding and correction, to raise or lower the social weight of the individual under discussion, irrespective of his or her biological sex. Gendered language was defined as language that rests on or refers to pre-fixed notions of gender roles, and stereotypes and standard images of masculinity and femininity and the relationship between the two. The investigation concentrated on metaphors that refer to activities or spheres that have traditionally been associated with one sex rather than the other as methods for analysing the way gender is used in the evaluation of scientific standards and excellence. It showed that feminine characteristics were used to feminise 'undesirable' male or female applicants and cast doubt upon their credibility as scientists. Positive qualifications were often expressed in more masculine language, the language of science, for instance, the expression 'to wrestle with a difficult task' is considered more positive than to 'have something knitting', which trivialises a reflection. From these results it appears that gender matters in the recruitment for professorships. However, this result was not confirmed by Lindberg et al. (2003) who found no systematic differences in non-neutral adjectives used by referees.

Gender between 'sameness' and 'difference'

Cultural interpretations of gender vary strongly – even within Europe. Are men and women essentially the same and is equality realised by equal treatment of men and women? Or are men and women essentially different from each other (whether by biological sex, or by historical traditions) and is equity realised by different treatment? Both positions can be interpreted as the extremes of a continuum; many scholars and policy-makers use a mixture of arguments depending on the context. Research on gender and science is clearly inspired by both visions, sometimes focusing on sameness (e.g. competences), sometimes on differences (e.g. gender differences in research issues).

Thorvaldsdóttir's paper provoked a discussion on what exactly is meant by 'masculine' and 'feminine' language, and the methodology that had been used, especially the independent criteria by which words were classified as masculine and feminine. This issue highlights one of the major methodological dilemmas within gender studies: analysing gender mechanisms without an implicit and unwanted reproduction of gender stereotypes. Moreover, gender-neutral concepts are very often used in a gendered way.

Social dynamics of scientific excellence: gender problems in practice

This chapter provides a short description of the social dynamics of research production and evaluation and the operation of gender in these processes. We can distinguish four stages:

1. setting the agenda for research
2. publications and citations
3. evaluation and assessment processes
4. transparency and accountability

Definition of research agenda

Science as a social institution has been dominated by men (white and socially privileged) and their views of what is important and relevant. We can expect this to have a strong influence on dominant discourses and established research agendas and paradigms. "*The scarcity of women in senior positions in science inevitably means that their individual and collective opinions are less likely to be voiced in policy- and decision-making processes.*" (She Figures 2003: 73). This also means that women are contributing less than men to shaping the big research questions of the moment. As long as women do not have a position in the power structure their influence in setting the research agenda is limited. If we want more women in the sciences, it is necessary to reflect on epistemologies and on research topics. This immediately leads to the question whether or not women have specific research interests. Addis (this volume) cites some evidence suggesting that men's and women's interests do not completely overlap. As regards economic sciences in Italy, on average women seem more interested in more 'socialised' fields, such as development economics, or in specific methodologies, such as narrative analysis. But this statement provoked a lively debate on the essentialist presuppositions. Sometimes, women do have different interests, as a result of differences in positions, historical traditions and roles, but not based on essentialist grounds arising from their sex.

> **More women in science, does it make a difference?**
>
> In many research fields there is no difference between the numbers of men and women if these are counted by research topic; in other fields, however, there is a difference. For instance, the Women's Health Initative in the USA: " … *higher numbers of women in power are thought to be a force for change in the research agenda"* (Cozzens, this volume). Women scientists identified that women were being left out of clinical studies, although the results were being used to diagnose and treat them. Addis found that women had more interest in the 'soft' issues of economics (developing countries) whereas men were also engaged in the hard finances and economic theories and models.

Women still constitute the marked category. Work which focuses explicitly on women, and do not mention men, immediately draws attention to itself (Griffin). As a rather new field, not well established at the institutional level, sometimes critical towards mainstream research and often interdisciplinary, gender studies is vulnerable to a negative judgment of its scientific quality. Gender-based research faces some difficulties in being acknowledged (Mottier and Griffin, this volume). This leads to a circular situation: rather limited publication opportunities, a lack of established positions and, as a result, low prestige. Moreover, the lack of established positions results in intense competition between gender scientists (Blagojevic, this volume). According to Rees (this volume), gender studies is vulnerable in the UK's RAE system as there is no separate panel for this interdisciplinary area, despite the fact that there is a sizeable body of work across a range of disciplines that could be described as falling into this area. "… *the RAE in effect (…) undermines an interdisciplinary field that is the focus for a considerable concentration of research by women*" (Rees, this volume).

In the evaluation of the scientific quality of research programmes, the incorporation of gender issues does not seem to be an important consideration. A review of existing research assessments of several scientific fields, carried out by Mottier (this volume), showed that many evaluators do not integrate gender into this work as an important category, even though some fields clearly have a gender dimension, such as pedagogics and management. In the Netherlands, research assessments within these fields were published in 2002 without – or with very little – reference to gender issues. The committees usually comprise male scholars and, even in the field of pedagogics, it seems as if excellence primarily refers to fundamental research: scholars engaged in social consultancy and policy advice performed badly in the assessment. Only one research assessment committee engaged women with gender expertise – socio-cultural sciences – and this influences the scope of the assessment. Gender issues are part of the evaluation: "*it is good that a gender dimension is fully integrated into most aspects of the programme rather than being a separate field of study*" (Assessment Report Socio-Cultural Sciences, in Mottier).

Social dynamics of publications, citations and success

It is not solely the search for excellence that is steering the selection process of publications and applications; other targets, such as the profile of the journal or the aim of the grant, also play a role. Some fields and issues are marginalised and therefore work in them is difficult to publish and/or get subsidised. In every organisation – an editorial board, an advisory committee, a research council – many stakeholders are represented. Visibility and scientific recognition are cornerstones for a successful selection process. It is not only talent and merit that decide whose papers will be published or whose application will be approved; this is also affected by social capital, especially in the current situation of overproduction of publications. In order to distinguish oneself as a scientist, publication alone is not enough; publications must be read, discussed and cited. "*Publications need personal representation within the scientific community*". Therefore, participation in academic networks is important and having the right (formal and informal) connections seems to be crucial in a successful career trajectory (Gupta et al.). Participation in specific networks is not only relevant in itself – the overall range and density of networks in which one is involved may also be significant for career opportunities. Having a wide range of 'weak ties' (Granovetter, 1974) or fewer 'structural holes' (Burt, 1992) is fruitful for finding out about opportunities, especially in contexts in which labour markets are less formalised.

Due to a lack of social capital – by which we mean access to resources and positions of power – women scientists run the risk of under-citation, the Matilda Effect mentioned above. Expanding on the statements of Gupta et al. (this volume), Feller suggests that, compared to men, women may experience disadvantages when having their work published, read, and thus cited, due to their lack of visibility and social capital. Scientists pay most attention to well-known or already established researchers. Citations may be by-products of participation in larger networks as well as measures of intrinsic scientific quality. Reputational status creates conditions for future success. In this respect, we can suppose that affiliation with the decision-makers for publication and research funding affects the opportunities of scientists. From this point of view, citations do not mirror 'quality' in an unambiguous way, but rather represent a mixture of 'quality' and 'social embeddedness'. A low citation score is not synonymous with a 'low quality of scientific work'. From this aspect, involvement in professional networks is an important strategy for improving the position of female academics.

However, as Blagojevic stated in her presentation, in countries in transition (such as some in Eastern Europe) networking does not mean the same as it does in Western European countries. "*In countries facing serious difficulties with transformation, nepotism not meritocracy prevails*". In that case, networking could be counter-productive for meritocracy. The ambivalence became very clear in the debate that followed the presentation of this paper. On the one hand, networking provides access to information and sometimes positions and, on the other, it is time-consuming and exclusionary.

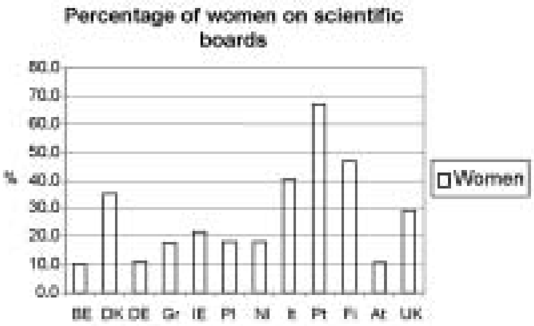
According to Valian (1998), the accumulation of advantage is an important explanation for the gender differences in scientific careers: success leads to more visibility and new successes, and enhanced reputation leads to more citations and more success in grant awarding and subsidies. This becomes an obstacle for newcomers in the field as they have to become visible and to compete with those who are already 'known' (Sandström and Hällsten, this volume). An instance of this is seen in a study by the Swedish Research Councils (MFR and NFR): most proposals were submitted by those seeking a renewal (circa 45%) and by people who were already known to the council as a result of former successful applications (circa 45%). Only a minority of the applicants were new to the research councils (10%). The first group was the most successful (85% received a subsidy) while newcomers were less successful (only 8%). This poses some questions about the often unintended effects of these systems regarding researchers who cannot rely on an established position.

Definition and selection of peer reviewers and composition of selection committees

Gate-keeping is one of the scientist's important functions. Gate-keepers are undoubtedly in a key position to influence the definition, evaluation and development of scientific excellence (Husu, this volume). Members of research councils, members of selection committees, peer reviewers, referees, editorial boards, are all in a position to define what is excellent, to distinguish between what is good and what is not. Gate-keeping can function as a process of control and exclusion, but it can also function as a facilitator providing resources, information, and opportunities. According to the ETAN report (Osborn et al., 2000), gate-keepers are generally middle-aged male academics. Women are clearly under-represented as gate-keepers. Their limited presence in the highest levels of academia is mirrored by their under-representation as peer reviewers, on editorial boards, research councils and in other decision-making positions. Many institutional rules block women's opportunities to take part in gate-keeping processes, as women are more often represented in low status positions. Even in countries like Finland, which has a rather high proportion of female professors (21% in 2002) and a 40% gender quota for all public bodies, the gate-keepers in the private research foundations are predominantly male (86% of the board members of the largest research funding foundations (Husu, this volume). The National Research Councils are approaching gender balance because they are government appointed in Finland and their composition has had to follow the quota paragraph of the Gender Equality Act since 1995. However, the proportion of women was clearly lower among referees who evaluate the funding proposals for the Research Councils. In 1999, only 16% of these were women and in 1999-2002 the proportion of women varied from 11-12% (in natural sciences and technology) to 33-43% (health sciences) (Husu, this volume).

Women in decision-making

The She Figures 2003 showed that the composition of scientific boards in most European countries is rather unbalanced. Only in Portugal do more women then men participate on the boards, while the Finnish boards approach 50%.



(Source: She Figures 2003: 76)

One of the important declared strategies to minimise gender bias in academic decision-making is a gender balanced composition of scientific boards.

Strong male dominance is also evident in other institutions in the power structure of the sciences: editorial boards, peer panels, and selection committees for professorships. The ways in which these organisations and groups are recruited are rather informal and lack transparency, the composition is often unbalanced and the male élite is over-represented. In many cases, new members are invited on the basis of their expertise or their networks, not on their ability to make reliable, unprejudiced judgments. Many criteria are marked by a kind of elasticity, and can be mobilised in order to make a positive or negative judgment. Moreover, issues like 'how to conduct assessments', 'how to use criteria' and 'how to avoid gender bias' are not recognised as important aspects influencing the professionalism and accuracy of selection and evaluation procedures (Mottier, this volume).

Processes: transparency and accountability

It has been suggested that, generally speaking, three main mechanisms are used to decide promotion: co-optation (discretionary promotion by those at the top, who are not accountable and do not state the criteria they use), selection (promotion by accountable people at the top, according to certain explicit criteria), and election (promotion by choice of the peers). Of these, selection appears to be the most transparent, while many appointments for professorships seem highly influenced by co-optation (Addis). Although more research is needed on this issue, many participants in the workshop support processes based on explicit standards, transparent procedures, and more public accountability as a means to aid screening for excellence, with positive spillover neither restricted to nor specific to gender issues. Evidence, however, shows that the opportunity loss resulting from the lack of transparency is highly gender specific: too much female talent is wasted.

Several experiments on gender-based double standards carried out by Foschi and her team were reported during the workshop. They clearly indicated a double standard in assessment processes. Different requirements were applied to men and women in assessing each other's competence. Moreover – and more importantly from a strategic point of view – this work showed that the effects of double standards decreased when the assessors were held accountable for the results by making the assessment public and known to the assessed. These experiments also showed that providing explicit standards rather than allowing assessors to generate and use their own criteria reduces the gender bias. Double standards flourish if assessments, assessors and criteria are not made public, leaving much room for subjective and uncontrolled judgments.

Indirect gender bias: crossing borders

The rather monolithic standard model of scientific excellence – with bibliometrics and peer review as the dominant measurement of scientific quality – does not fully coincide with the heterogeneity of what actually happens in the sciences. As Griffin stated in her paper, "*There is, as yet, in the literature concerned little recognition of the fact that assessment of scientific excellence, far from being an exercise in disinvested and disinterested judgments, is one of situated knowledge-making, reproducing the cultures from which it emanates.*" (Griffin, this volume: p. 10) Four dimensions of 'situated decision-making' were discussed during the workshop: disciplinary differences, mono- and interdisciplinarity, different modes of science, and geographical location at the centre and periphery. Sometimes the relationship with gender is quite clear – e.g. the disciplinary differences – but in some cases the gendered character remains indefinite. There is – for instance – some evidence that women scientists have a stronger tendency towards interdisciplinary research and towards research aiming at social issues, but the results are not unequivocal.

Diversity among disciplines and research fields

Measuring scientific quality using bibliometrics is clearly anchored in publication traditions of the natural and related sciences that are generally well organised at an international level, with publication occurring in scientific journals (e.g. psychology). These standards reflect the normative ideal of the natural sciences and have come to represent a model for all scientific research. This seems to cause some problems for measuring scientific quality in the humanities and in many social sciences (Mottier, this volume). The coverage of journals by the 'citation index' is much lower in the latter than in the natural sciences and, moreover, in many disciplines scholars prefer to publish books instead of articles in journals. In general, the normative ideal of 'real science' is more or less identical with natural science, and recourses and infrastructures for research are unevenly distributed among the disciplines. Since women are proportionately more actively engaged in the social sciences and humanities, this results in an unintended but nonetheless active gender bias that affects the opportunities for the scientific careers of female scholars and their relative success rates.

Monodisciplinary versus interdisciplinary fields

Participants in the workshop agreed that in the last few years there has been growing recognition of the need for interdisciplinary research – but this recognition is only apparent at the level of discourse. However, many assessment systems are rooted in monodisciplinary activities and, in fact, assessment regimes produce or reproduce disciplinary boundaries and elicit disciplinary behaviour. Interdisciplinary fields do not belong to the core issues or the core community of established fields; they are not well organised within research councils, and remain largely invisible. Similar processes take place at the individual level: interdisciplinary or multidisciplinary activities are often invisible. Publications in high quality journals from other disciplines have almost no impact on the personal network or the scientific reputation, as Feller stated during the workshop. In most cases, such publications are simply not recognised or acknowledged. Although interdisciplinarity is definitely part of the discourse, the system seems to operate at the disciplinary level.

This need for interdisciplinary research is closely related to the next issue: the missions and tasks of universities and scientists in a knowledge-based society.

Different modes of science

One of the central lines of debate during the workshop referred to the changing position of the sciences as a social institution in the emerging knowledge-based society. Different concepts have been developed to cover these transformations. Gibbons et al. (1994) describe them as Mode 1 versus Mode 2 models of knowledge production and transmission; Laredo (this volume) addressed it as a 'third mission' of the universities – a responsibility to ensure more direct links with societal and economic needs. Brouns (this volume) refers to it using the metaphor of Mount Olympus versus the Agora. The classic but still powerful metaphor for science is the Olympus model which situates scientists, in their unselfish and disinterested quest for truth, at the top of the pyramid, far removed from the concerns of everyday life. In the Agora model, science is analysed as a societal practice, tightly bound to other such practices. In the context of the knowledge-based society, the sciences are moving into the Agora, but this is hardly recognised in the evaluation systems. Scientific knowledge refers to creators, transmitters and users (Blagojevic, this volume), but only the first ones are acknowledged in the dominant system of measuring scientific quality.

The *Centre de Sociologie de l'Innovation* has designed a 'research compass card model' to map the complex world in which researchers have to operate (Laredo, this volume). It would be counter-productive for research institutions to consider only academic excellence when, at the same time, policies demand other missions to be developed. What is needed is for indicators to be developed for performance in other research activities. Laredo argues for a transition from scientific excellence to scientific performance, which clearly engages in scientific activities in a broader scope. A wide range of achievements is needed to be successful as a scientific group, varying from designing research, addressing several audiences, acquiring research money and negotiating with governments and enterprises to gaining new resources and facilities for research. This wide range of tasks is not within the reach of

Enwise expert group on women scientists

In February 2004, the Enwise expert group on women scientists presented a report to the European Commission. This report investigates the situation of women scientists in the Enwise countries: Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia. The report highlights the influence of specific gender policy implemented in these countries during the communist regime. The proportion of women among researchers in Enwise countries is higher than in the current 15 EU Member States. Analysis reveals that even in countries where the presence of men and women is fairly balanced, there are gender differences in the concentrations across the fields of sciences, whereby women are absorbed in low-expenditure systems. It shows that although formal legislation is an absolutely necessary condition, it is not sufficient to guarantee equality. The consequences of the transition have left women scientists in a more vulnerable position.

individual scientists, but presupposes a collective effort from research teams. But how do we manage assets when taking the collective dimension of research into account?

Universalism versus national/regional research traditions

According to Blagojevic, some research valued as excellent in the context of production does not fit in the Anglo-American traditions and journals and is thus rejected. This is clearly a case of situated decision-making: what 'good science' is depends on the context of evaluation and the inherent evaluation systems. The dominance of Anglo-Saxon traditions might cause problems for researchers in countries at the semi-periphery. First of all, they face some complex problems related to the epistemic culture in their own countries, which may be more problem-oriented than knowledge-oriented, and secondly, they are confronted with an asymmetry in the roles of Eastern and Western scientists. As Schiebinger stated some years ago: "*the Science Citation Index rarely includes sources in languages other than English, so the hair-splitting measures of productivity and influence do little to consider science in a global context*" (Schiebinger, 1999: 50).

Many concepts used in research are not universal, they are context-specific, as Blagojevic argued. This means that in semi-periphery countries there are circles of translation that make the scientific discourses accessible to researchers in these regions. Adaptation to international conceptual frameworks is necessary in order to get published. However, this might cause some tension since these conceptual frameworks are not always adequate or relevant: "*some theories are irrelevant to our experiences*" (Blagojevic).

In conclusion, many of these elements in the actual functioning of the system cause an unintentional tendency towards conservatism. The system of evaluating scientific quality has become much more than an accounting device and is being taken much more seriously. In fact, in many situations it is decisive for new appointments, distribution of resources, and publication opportunities. Scientists react accordingly, trying to establish good prospects. It is necessary to take into account the performative effects of these kinds of evaluation systems – i.e. the fact that the scientific product will often tend to change as a result of the evaluation system as scientists try to be evaluated positively. According to Feller (this volume), the emphasis on bibliometrics is too great because even an unbiased measure, when applied in a biased system, will produce biased results. Therefore, too rigid a use of bibliometrics may obstruct a refined measurement of achievements and contributions of women scientists.

Instead of adapting traditions and behaviour to the standards, it is argued that it is necessary to broaden the spectrum of activities and achievements to be included in the definition of scientific excellence. It is important to include other dimensions of scholarly practice, such as education, participation in committees, administrative tasks, external consultancy, and contribution to public debates. In other words, it is important to emphasise not only production, but also relevance and the different users of scientific knowledge. Measures are meaningful if they are based on the context of their production and in the uses of the knowledge. But the question of how exactly this science-in-context is measured and what kind of criteria should be applied remains subject to debate.

Conclusions

During the workshop, the pursuit of excellence was described by many participants as *a set of practices framed within a 'discourse'* used by the scientific community to organise its self-governance. The workshop addressed the problematic aspects of some contemporary systems in general, as well as in relation to gender. The workshop frequently discussed the issue of the validity of existing definitions and methods for measuring scientific excellence, in connection with the variety of scientific practices, interdisciplinary research, and mainstream versus periphery in terms of issues, methods and location.

The specific gender dimension appears as the continuation, in the top tiers of science making, of the same mechanisms that prevent more women from participating in larger numbers, even at the rank-and-file level of scientific enterprise. Achieving the best results in terms of scientific productivity requires that the discourse on 'excellence' be reframed in such a way as to include all scientists, independent of their biological sex. The presentations in the workshop showed that the environment in which scientific inquiry takes place is not yet free of bias against women, and that this is linked to the subtle persistence of gender stereotypes. Persistence of these stereotypes generates many 'double standards': women are repeatedly requested to prove themselves worthy, as stricter criteria to assess competence are used in judging them. This may prevent some of the best scientific minds from growing to their full potential, and some innovative and very productive scientists from participating fully in scientific governance. It also creates persistent stress for those female scientists who keep working, often achieving excellent results.

That the scientific environment should be free of gender stereotypes is a statement that would most likely be agreed upon by the vast majority of men and women involved in science. Nonetheless, translating this statement into a strategy that really leads to a 'gender-equal' or 'gender-sensitive' scientific environment, one that is not gender discriminatory and that is capable of containing all genders and valuing all genders' abilities equally, has proved more problematic. The first step in this direction is cultural; it is the understanding, by the entire community, that the problem is there and must be addressed.

Several aspects of possible gender bias in the production and evaluation of scientific excellence were discussed in the workshop. Gender bias can occur (1) in the characterisation of scientific excellence, (2) in the criteria used to assess it, (3) in the choice of the explicit and implicit indicators for scientific excellence, (4) in the way the criteria are applied to men and women, (5) in the failure to integrate women in scientific networks, and (6) in the procedures through which criteria are applied to people.

(1) Characterisation of scientific excellence

The discourse about what characterises excellence is generally not subject to scientific evaluation, and the actual practices in each branch of science are often quite idiosyncratic. It is assumed that the scientist in each field somehow acquires, from his or her environment, a notion of what excellence is, and that there is no need for a critical evaluation of the concept and of its correspondence with actual practices. There is a dominant characterisation, the one adopted in the natural sciences, which produced standards that have been easy to adopt more generally but have been criticised as rather limited and prone to reductionism. Social sciences have adopted similar standards, introducing them as grounded in the practice of the natural sciences, but they do not acknowledge the diversity in missions and achievements, such as the so-called third mission of universities (link with society). Other scientific practices, in particular those from disciplines most popular among women, such as humanities and social sciences, are not adequately included in the characterisation, and the standards do not cover what excellence there might be in these disciplines sufficiently.

(2) Criteria

Criteria used to assess excellence are often described at a very abstract level (researchers need to be innovative, productive, coherent, etc.). These abstract nouns become frameworks for interpretations in a specific context or in epistemic cultures. Teaching and professional activities are usually undervalued, and if women have a systematic overload of these activities as a result of their temporal contracts and positions, this may introduce a source of gender bias. As underlined by Griffin, the boundaries within which criteria are applied, i.e. the definition of what falls within and what falls outside a given discipline, may be a very important source of bias.

(3) Indicators

Scientific excellence is not easy to grasp. The dominant image of the excellent scientist is more or less grounded in a male career pattern with an absolute dedication to science. Many people – especially those with family responsibilities – find it hard to live up to this image. Good indicators of excellence should take into account the fact that time-poor people can still be excellent scientists, and should measure scientific output not only in cumulative terms, but also in relation to the amount of time devoted to research. Many indicators explicitly or implicitly used to assess one's qualities do not accurately measure one's scientific products. Some of these indicators can be an obstacle for women, such as affiliation to established scientists who have access to many resources, position within a social network, and competitive style. Indicators that are unbiased per se (counting of publications and citations) may still produce biased results when applied to a biased system (e.g. if men cite men more often than they cite women).

(4) Assessment of competence based on application of these criteria

Many investigations presented during the workshop produced evidence of the fact that standards adopted for the attribution of scientific competence to women differ from the standards used when assessing men. In experimental settings, similar achievements led to different assessments of the task competence of men and women (Foschi). Gender-biased judgments appear to be pervasive: people of both sexes may apply double standards when working as evaluators of themselves as well as of others. Accountability seems a major strategy to reduce double standards. Evaluators who have to explain their judgment are more sensitive to the equal assessment of men and women. This means that transparent procedures may contribute to a gender-equal atmosphere.

(5) Networks

The workshop stressed the power of networks, emphasising their role in information exchange, agenda setting, control, and implicit decision-making. Therefore, it is important to increase women's participation in networks, which is also a *sine qua non* to be known as a scientist. Because networks are often quite informal and generated by pure co-optation, it is difficult to devise a clear strategy to increase women's participation. Moreover, we have to register some ambivalence towards suggesting that women invest more in networking because it is very time consuming. Although networks as exchange media are somewhat unavoidable and therefore are probably here to stay, a 'gender-equal' or 'gender-sensitive' scientific environment requires networks that are inclusive of women, and a shift of decision-making power away from male exclusivity and informal networks and into more formal and professional mechanisms.

(6) Procedures

Excellent scientists are those who have proved themselves through a set of different screening procedures. 'Hiring' procedures are different from 'funding' procedures which, in turn, are different from 'review for publication' procedures. The workshop underlined the need to think, in a more general, systematic and scientific way, about various procedures that are free from gender bias. Sources of problems were indicated in a) choosing the 'gate-keepers' in charge of a procedure, b) gender balance of panels in charge of procedures, c) transparency of the procedures, which includes advertising positions and criteria being made explicit, and d) accountability of the 'gate-keepers'.

Recommendations

Although individual participants may favour one solution over another, some consensus over lines of action did emerge:

Evidence

More research is needed for a better understanding of where the bias comes from. Gender is a deep cultural construct that operates at symbolic and institutional levels, and gender bias may assume different forms in different cultures, which may go unnoticed unless systematically explored. There are fields where women have fared better than in others; this needs systematic investigation.

General recommendation: Funding of research in some neglected areas, such as differences between disciplines, epistemic cultures, and national and regional contexts. These differences should be compared and investigated in order to improve our understanding of the gender dimension of science and scientific organisations.

i) *Assessment of scientific excellence: micro level*
   - Research on the attribution of scientific competence to men and women with highly comparable academic records, focusing on systematic research on key factors that affect the intensity and relative weight of a biasing practice (e.g. transparency of procedures, degree of accountability, disciplinary sex-category association, sex-category composition of assessors and assessed, national contexts).
   - Investigation of personal gender bias of assessors, especially the extent to which sex-category is viewed as an implicit indicator of scientific competence.
   - Systematic research on perceptions of excellent scientist in different disciplinary, organisational (universities, research institutes) and national contexts, compared to perceptions of women scientists and men scientists in the same context. Also investigate possible differences between men and women as regards these perceptions.
   - Analysis of performative effects of the standards of scientific excellence and the way male and female scientists adopt and identify with these notions in their own scientific practice.

ii) *Assessment of scientific excellence: institutional level*
   - Comparative investigation of procedures and protocols at universities and other research institutions and their effect on the gender dimension. Comparison should be focused on gender blindness, gender sensitivity and gender relevance.
   - Research on decision-making procedures of funding agencies and the genderedness of 'epistemic cultures' in comparative perspectives in different disciplines, including areas where women are relatively well represented.
   - Examination of constructions of masculinities within evaluations of scientific excellence and academic merit.
   - Investigation of gate-keeping in relation to research funding, recruitment and promotion in universities at the individual and institutional levels, their practices and policies.
   - Comparative analysis and gender impact assessment of initiatives taken in various institutional settings for measuring scientific excellence in a gender sensitive way or for defining transparent and fair methods for assessing merit, quality and productivity.
   - Comparative analysis of how women's studies and gender studies, as highly feminised interdisciplinary areas of research, are treated in the research assessment exercise and on the impact of these assessments on the field developments.
   - Elicit constructs of 'excellence' from a substantial sample of men and women in order to design a standardised list of criteria for excellence at a more operational level.
   - Elicit constructs of effective leaders in sciences from a sample of females and males, and content analysis.

iii) *Social dynamics of scientific excellence*

- Research on the impact of networking in science, both as an existing practice and as a strategy for gender equality. Examining gender and the inclusion or exclusion of women from Networks of Excellence would be a very relevant case study.
- Investigations of the relationship between gendered culture, practices considering scientific excellence, and career tracks of men and women scientists.
- Research on gender differences as regards the support – logistic, technological and financial – provided by the institutions to men and women scientists.
- Research on the relationship between mobility and the construction of scientific excellence.
- Investigation of family-related aspects, such as part-time work, and its effects on scientific excellence and the attribution of scientific competence.

iv) *Science and society*

- Investigation of the relationship between knowledge and policy-making; comparison of different European regions.

Awareness

An important first step is to make all scientists, male and female, aware of the extent and the consequences of the problem of gender bias in measuring excellence. In particular, those in charge of screening procedures should be trained to understand gender bias and its consequences, so as to minimise it.

General recommendation: Special training programmes on gender awareness, designed by gender specialists. Development of reading material on gender bias in evaluating research.

Field boundaries

Particular attention should be paid to the fact that the definition of the boundaries of a field of inquiry and the specific activity profiles of scientific institutions matter when deciding who is excellent.

General recommendation: Much greater recognition of the positive contribution of interdisciplinarity, new research fields and gender studies to scientific excellence. Discuss also the relevance of productivity measures as a primary way of assessing researchers' performances.

i) *Interdisciplinarity*. How to evaluate contributions that 'fall through the cracks' between two fields of inquiry? Working at the borders of a discipline creates some risk and this, in particular, is true of gender studies. An evaluation of interdisciplinarity, in contrast to the lip-service it usually gets, may be helpful.

ii) *Novelty of fields*. Creating a new field is often a matter of redefining the boundaries of an old field, so as to carve out the study of a specific terrain, with its own methodology, network, and shared knowledge.

iii) *Activity profiles.* Discuss the relevance of productivity measures, especially if other activities (teaching, management) are crucial to the actual functioning of researchers.

iii) *Gender studies.* In recent years, women scientists trained in many different fields have chosen to focus on studying gender from the perspective of their discipline. Gender issues have been explored along a spectrum ranging from the visual arts to literature, philosophy, psychology, history of thought, law, sociology, economics, and biology. There is definitely a field of gender studies, but it is difficult a) to define excellence within gender studies given that people come from so many different perspectives, and b) to make the contributions on gender count towards assessing performance in the original field of the gender scientist. From a purely career-oriented viewpoint, contributions to gender studies are devalued with respect to advancement in the original field – which has not prevented gender studies from growing.

Networks

The issue of networks is strongly linked to the issue of awareness and power. Scientists should be more aware of gendered difficulties in engaging in dialogue and networking with scientists of the other sex. More formal processes and criteria need to be set up to allow more women to pass the 'fuzzy screening' of networks.

General recommendation: gender balance in officially funded networks to be achieved by the imposition of recommended quorums formulated with reference to women's and men's respective presence in the field.

Procedures

Procedures for assessing excellence are not obvious or natural; there needs to be a critical examination of their interaction with gender.

General recommendation: in order to minimise gender bias, it is of particular importance that screening procedures be made more transparent and evaluations more public.

i) *Public advertising of positions*, thus helping to reduce the influence of informal networks that often exclude women.

ii) *Explicit standards of promotion or appointment.* The existence of explicit standards that apply to a large set of applicants tends to favour objectivity. On the other hand, a caveat is needed against the practice of tailoring the standards for a call or a position to favour one particular applicant.

iii) *Standards based on traits that are relevant* to the pursuit of scientific knowledge.

iv) *Use of appropriate indicators of performance*, rather than intransparent judgements.

v) *Indicators chosen to fit the life-cycle productivity* of scientists, both women and men.

vi) *Indicators chosen to fit the different scientific interests* of scientists, both women and men.

vii) *Public availability of the records* of a procedure – even making proceedings public to a certain extent.

viii) *Accountability* of the members of panels in charge of procedures.

ix) *Composition of panels balanced in terms of gender ratios*, at least with respect to the promotion of women among applicants or senior experts in the field. Particular attention should be paid to the selection criteria for editorial board members and referees of scientific journals.

x) *Gender audits* on women's participation as well as on the gender dimension of research programmes should have effective consequences for a mark in evaluations.

xi) *'Professionalisation'* of selection procedures by training panel members, chairs and members of editorial boards to look for and to correct gender bias.

xii) *Recognition of that change – to reduce gender bias – involves all participants in the process, both men and women, not only women.* There should be excellence in practices of definition and measurement of scientific excellence, and not only in excellence in science itself.

General recommendation: as procedures are different in different countries, the above requirements regarding transparency need to be adjusted to fit each national situation, in a coordinated effort at the European Union level. In addition, research-funding procedures employed by the EU should be revised to ensure transparency.

NOTES

(1) Investigations presented during the workshop show that the Wennerås and Wold article has had a notable influence on the success rates of female applicants in Sweden, which have increased considerably due to what Sandström calls 'The Wold Effect'. Since the publication of the Wennerås and Wold article, Swedish research councils – and many others, supposedly – have become more aware of the possibility of gender bias in the peer review of research grant applications.

(2) Henneberg (1997) argues that evaluators sitting on grant committees tend to be selected because they are working in a problem area similar to that a grant application proposes to solve. However, since scientists working in similar areas tend to know each other and either dislike each other and compete amongst themselves or like each other and collaborate, the result of a review of grant applications depends on political and financial elements external to the application.

REFERENCES

Bordons, M., Morillo, F., Fernandez, M.T. and Gomez, I. (2003) 'One step further in the production of bibliometrics indicator at the micro level: Differences by gender and professional category of scientists', *Scientometrics*, 57: 159-173.

Burt, R. (1992) Structural Holes: *The Social Structure of Competition*, Cambridge, Ma: Harvard University Press

Cole, J.R. and Cole, S. (1985) 'Experts', 'consensus' and decision-making at the National Science Foundation', in K. Warren (ed.) *Selectivity in Information System- Survival of the Fittest.* New York: Praeger, 27-63.

Essed, Ph. and Golberg, D. (2002) 'Cloning cultures: the social injustices of sameness', *Ethnic and Racial Studies*, 25: 1066-1082.

European Commission (2003) *She Figures 2003 – Women and Science Statistics and Indicators*. Luxembourg: Office for Official Publications of the European Communities.

European Commission (2004) *Waste of Talents: Turning Private Struggles into a Public Issue. Women and Science in the Enwise countries*. Brussels: European Commission.

Granovetter, M.S. (1974) *Getting a Job, A Study of Contacts and Careers*, Cambridge, MA: Harvard University Press.

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

Henneberg, M. (1997) 'Peer review: the Holy Office of modern science', *Natural Science*, Article 2.

http://naturalscience.com/ns/articles/01-02/ns_mh.html

Kanter, R.M. (1977) *Men and Women of the Corporation*. New York: Basic Books.

Knorr Cetina, K. (1999) *Epistemic Cultures How the Sciences make Knowledge*. Cambridge, Ma: Harvard University Press.

Lindberg, L., Riis, U. and Wallin, C. (2003) *Gender Equality in Swedish Higher Education – A Complex Issue II*. Paper presented during ECER conference, Hamburg.

Long, J.S. (1992) 'Measures of sex differences in scientific productivity', *Social Forces,* 71: 159-178.

Merton, R.K. (1942) *The Sociology of Science*: Theoretical and Empirical Investigations. Chicago and London: The University of Chicago Press.

Merton, R.K. (1968) 'The Matthew Effect in science', *Science, 56-63.*

MIT *(Massachusetts Institute for Technology) (1999) A Study on the Status of Women Faculty in Science* at MIT. Women Faculty of Science. Cambridge, Ma: MIT.

Nilsson, J. (1997) 'Utvärdering av de vetenskapliga rapporterna (Evaluation of the scientific reports)', *MFR Informerar* (Information from the Swedish Medical Research Council), 3: 8-9.

Osborn M., Rees, T., Bosch, M., Ebeling, H., Hermann, C., Hilden, J., McLaren, A., Palomba, R., Peltonen, L., Vela, C., Weis, D. and Wold, A. (2000) *Science Policies in the EU: Promoting Excellence through Mainstreaming Gender Equality*, ETAN Report. Luxembourg: Office of the Official publication of the European Communities.

Rossiter, M. (1993) 'The Matilda Effect in science', *Social Studies of Science*, 23: 325-41.

Sax, L.J., Hagedorn, L.S., Arredondo, M. and Dicrisi Iii, F.A. (2002) 'Faculty research productivity: Exploring the role of gender and family-related factors', *Research in Higher Education*, 43: 423-446.

Schiebinger, L. (1999) *Has Feminism Changed Science?* Cambridge, Ma: Harvard University Press.

Sonnert, G. and Holton, G. (1996a) 'Career patterns of women and men in the sciences', *American Scientist*, 84: 67-71.

Sonnert, G. and Holton, G. (1996b) *Who Succeeds in Science? The Gender Dimension.* New Brunswick, NJ: Rutgers University Press.

Stamhuis, I. H. (1995) 'A female contribution to early genetics: Tine Tammes and Mendel's Laws for continuous characters', *Journal of the History of Biology*, 28: 495-531.

Valian, V. (1998) *Why so Slow? The Advancement of Women*. Cambridge, Ma: MIT.

Wennerås, C. and Wold, A. (1997) 'Nepotism and sexism in peer-review', *Nature,* 22: 341-343.

Zuckerman, H. (1987) 'Persistence and change in the careers of men and women scientists and engineers: A review of current research', in L.S. Dixon (ed.) *Women: Their Under-Representation and Career Differentials in Science and Engineering*. Washington: National Technical Information Service, 123-156.