# Scientific EVALUATION in Korea

*WON GEUN RHIE

*The task of evaluating science and technology is becoming more important in every country trying to develop and improve overall national R&D capability in an effort to sustain national prosperity and public welfare. However, the way in which evaluation organisation is approached in each country may differ from nation to nation reflecting national R&D systems and S&T policies as well as the specific roles of scientists and policy-makers at various levels.*

*Over the past two decades or so, Korean evaluation system has been transplanted from the West and they modified. In addition, Korea borrowed elements from the Japanese. Whether such systems have been well rooted in our specific circumstances is a critical question.*

*In this context, it would be meaningful to diagnoses the current Korean evaluation system and also makes some suggestions for improving this system. For this purpose, this paper also attempts to approach to the theoretical aspects of evaluation as a starting-point.*

## 1. Some evaluation definitions and concepts

Evaluation emerged from the general acceptance of the scientific method as a means of dealing with social problems. However, despite historical roots that extend to the seventeenth century, the widespread employment of systematic, data-based evaluations is a relatively modern development (Rossi and Freeman, 1989). Evaluation which was first initiated in the fields of education, public health, and development projects such as rural and urban community development became commonplace in all fields with the rise of public policies and the increase in the need for government intervention and the development of more complex methods of applied social research. A significant number of social scientists had been advocating the application of rigorous social research methods to the assessment of community action programmes, and evaluations were implemented more frequently (Freeman, 1977).

* Assistant Head, The Academic Exchange Team of Korea Science and Engineering Foundation (KOSEF)
* PhD student, PREST, The University of Manchester in the U.K
* 한국과학재단 학술조성팀 선임행정원 이원근 (영국 만체스터대학교 과학기술정책과 박사과정)

Evaluation in science and technology, a relative latecomer to public policy, was initiated only to estimate the quality of research within the limited scientific boundary, with little interchange with communities in its earlier stages. However, as science and technology has become a more important factor in economic and social development, the evaluation of science and technology has been recognised as a crucial input to policy-making at the national level as well as a tool for systematically estimating research performance at an individual and institutional level.

***Why do we evaluate?***

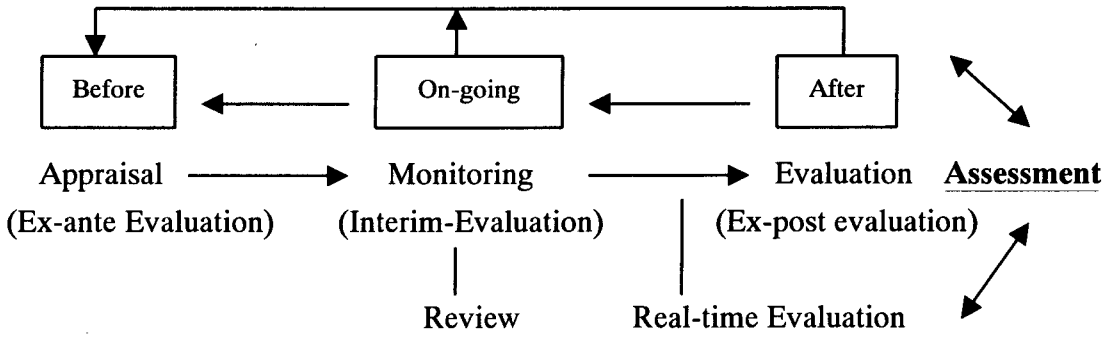
In every country, evaluation is now a common and important interest. Evaluations may be undertaken for a variety of reasons. In practical terms, evaluation is carried out in order to systematically collect useful information about the activities, characteristics, and outcomes of programmes, to make judgements about the programmes, to improve programme effectiveness, and/or inform decisions about future programming (Patton, 1986). Evaluation is designed to generate specific suggestions for future policy, reaching a view on value for money (OECD, 1990), and to provide the basis for informed decision making at all levels of management in the initiation, selection, direction and termination of R&D (Cabinet Office, 1989).

Evaluation, both as an academic discipline and as part of the policy environment, has taken place against a background of declining or even negative rates of growth in the resources available for science and technology. At the same time, the expansion in new technologies has focused attention upon the exploitability of science and technology and has increasingly led science managers to use this argument to justify their resources. Inevitably, competition for resources has intensified both within and outside the science community. Externally, value-for-money has to be demonstrated both in terms of efficiency and of impact so that science and technology can compete with other demands upon public funds. Internally, demand for resources has outstripped supply and has made selectivity and concentration on the only viable strategies (Georghiou, 1989).

***What is the evaluation?***

In a narrow sense, it is merely the process of checking, at the completion of R&D work, to what extent and how the objectives were met. However, it is also used as a conceptual tool for planning, formulating and drawing conclusions about the entire R&D activity in its loop process as a whole. There are different concepts or definitions of the activities involved in the process of evaluation.

**Circular process of evaluation**




***Appraisal**, the process of identifying required benefits, defining rationale, setting objectives, examining options and weighting up the costs and benefits in the initiation. This is also called ex-ante evaluation

***Monitoring**, regular checking of progress against plan, being a systematic tool for identifying strengths and weakness and providing information to improve R&D work (Interim-evaluation)

***Review**, a specific and formal exercise carried out by people independent of the line management, to take a closer look and to produce a set of recommendations regarding the future of the work

***Evaluation**, a circular process, which helps proposed scientific activities to be planned, formulated and achieved successfully together with appraisal, monitoring and review process. Also known as Ex-post evaluation

***Real-time Evaluation**, continuous evaluation so that the lessons could be acted upon immediately as the R&D activities unfolded.

***Assessment**, a general term covering the integrated activity of appraisal, monitoring and evaluation as a whole.

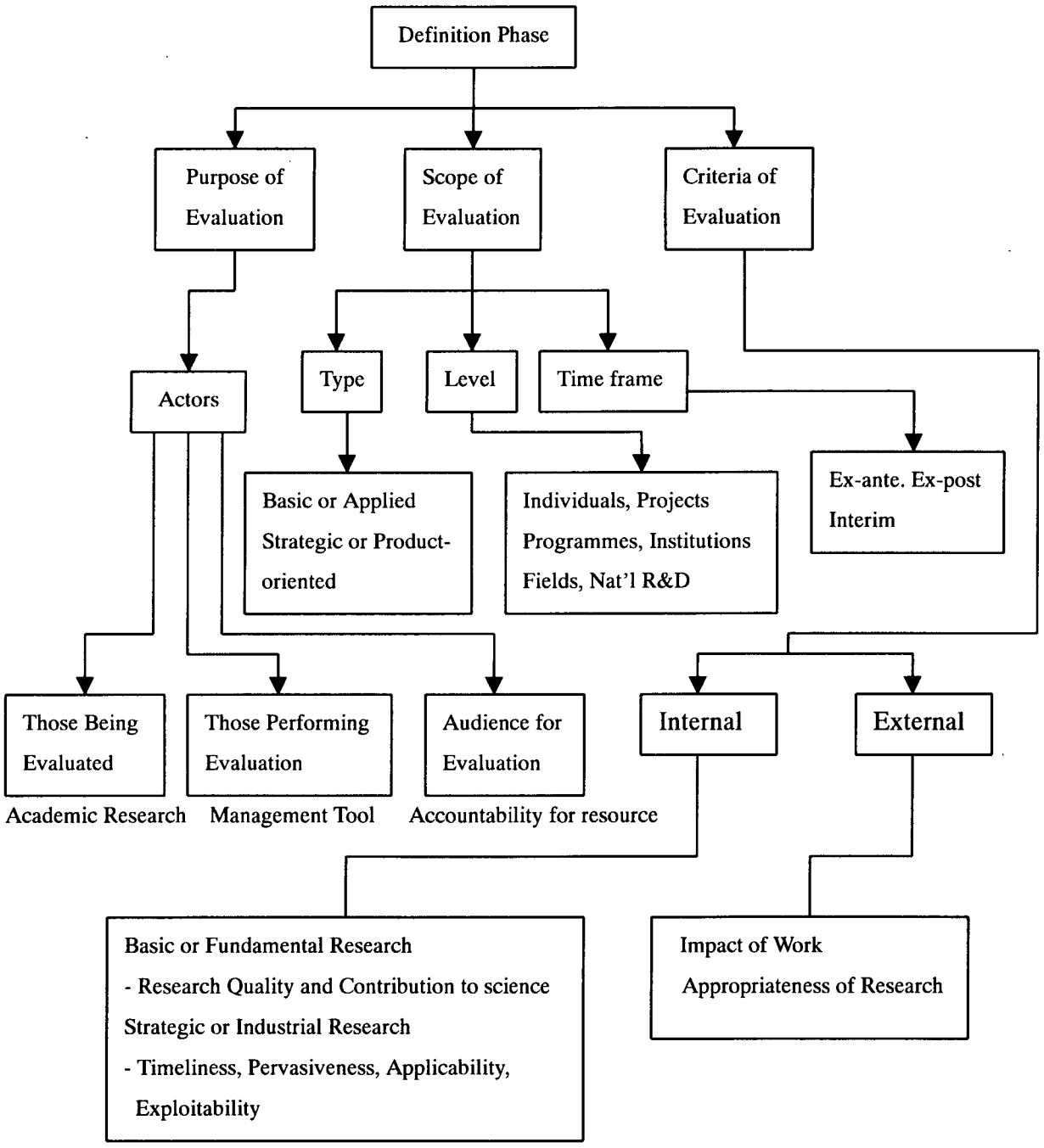
## 2. Conditions for an evaluation

Before starting an evaluation exercise, in order to gain the most useful evaluation outcome, it is first necessary to set up some specific target guidelines, such as what is to be evaluated, by whom, by which criteria, for whom and to what purpose.

In the initial stage of an evaluation, the task ahead needs to be defined in terms of the scope, purpose, criteria and the organisation of the evaluation process itself. The scope of an evaluation addresses the type of research to be evaluated, its boundaries and the timing of the evaluation exercise. The criteria provide the basis for the issues to be investigated. The purpose needs to be clear to all concerned and the organisation involves selection of the evaluators, the operating procedures and decisions on the level of available resources. The critical element is the interface between the evaluation and the wider policy making process (Georghiou, 1989).

These conditions must be settled prior to and not during the study and be clearly articulated beforehand. These elements are interrelated in carrying out the evaluation. In an effective evaluation, it is obvious that some elements may need to be adjusted. Once the relationship between these elements has been firmly established, the way is open to choose the appropriate methods and techniques to achieve the objectives of the evaluation (OECD, 1987).

**Organisation of evaluation**



## 3. Evaluation methods and their characteristics

Once such evaluation conditions as scope, purpose, criteria and organisation are established, the next steps are to choose the appropriate methods and relevant techniques to achieve the objectives of the evaluation (OECD, 1987). The evaluation method chosen varies from country to country and organisation to organisation, according to where the emphasis will be placed. However, in broader terms, the evaluation methods used are largely classified by reference to peer review, the qualitative method, and bibliometrics (publication etc) the quantitative method. Peer review is currently the predominant method being employed to evaluate research and aims at a consensus of expert judgement. In order to gain more useful information regarding the objects or targets evaluated and stand a reasonable chance of a sound judgement being reached by the experts, these two kinds of methods must complement each of them from the start.

### *Peer review as the qualitative evaluation method*

The use of experts representation of the scientific community, i.e. peer review, is the most traditional and also most important way of evaluating quality in science. The first historical use of peer review occurred shortly after the founding of the Philosophical Transactions of the Royal Society in 1665 (Zuckerman & Merton, 1973). In authorising its publication, the Royal Society stipulated that all articles to be published should be reviewed first by some members of the same field. Thus was born the gatekeeping function of scientists passing judgement on the quality of other scientists' work. Over the centuries, this practice has evolved into the system of journal refereeing, and has branched into the use of peer review as a mechanism for the allocation of research support (Chubin and Hackett, 1990). Currently, peer review is used by nearly all research funding agencies of the world in their evaluation procedures.

*What is Peer Review?* There are many synonyms of peer review such as peer advice, peer evaluation, peer judgement, quality control, peer censorship, merit review, refereeing, and so forth. Although these terms may differ depending on the particular situation, generally, peer review can be understood as a generic term encompassing all of them.

Peer review in science is an organised method for judging the scientific merit, intellectual excellence or importance of a piece of work, and is used by one or more individuals supposed to be sufficiently knowledgeable in the same or relevant field to certify the correctness of procedures, establish the plausibility of results, and allocate and distribute scarce resources such as journal space, research funds, recognition, and special honours in effective ways. (Chubin and Hackett, 1990; Westerheijden, 1991; OECD, 1987; Kruytbosch, 1989; Boden

report, 1990)

Bozeman (1993) classified peer review into three categories according to the level of its impacts on decision-making: pre-emptive, traditional, and ancillary peer review. Chubin and Jasanoff (1985) also classified it into three categories by whether its application was the review of articles to be published in journals, the selection of research proposals or the evaluation and interpretation of research results for decision-making.

Peer review should function as an effective mechanism for allocating resources and communicating priorities to scientists and should promote accountability in science. It is also expected to be responsive, providing a mechanism for policymakers to direct scientific effort and for new science to shape future research. In addition, it should be rational in the sense that its inner workings should be knowable and seem reasonable to the participants; and fair, adhering to societal norms of equitable treatment as well as scientific norms of universalism and disinterestedness. Finally, insofar as peer review measures scientific performance, it should adhere to technical standards of good measurement such as validity and reliability. However, the peer review system has been criticised on several counts.

Firstly, peer partiality is an increasing problem as research facilities become concentrated in fewer, larger centres making peers with no vested interest in the review process increasingly difficult to find.

Secondly, the 'old boy network' often results in older, entrenched fields receiving greater recognition than new, emerging areas of research, while declining fields may be protected out of a sense of loyalty to colleagues. Thus, the peer review process may often be ineffective as a mechanism for restructuring scientific activity.

Thirdly, the halo effect may result in a greater likelihood of funding for more visible scientists and for higher status departments or institutions.

Fourthly, reviewers often have quite different ideas about what aspects of the research they are assessing, what criteria they should use and how these should be interpreted. The review itself may vary from a brief assessment by post to site visits by panels of reviewers.
Fifthly, the peer review process assumes that a high level of agreement exists among scientists about what constitutes good quality work, who is doing it and where promising lines of enquiry lie, an assumption that may not hold in newer specialities.

Lastly, resource input to the review process, both in terms of administrative costs and of scientists' time, are considerable but usually ignored. (King, 1987 ; Chubin and Hackett,

1990 ; Kostoff, 1994)

Although there are some negative aspects to peer review, ABRC (1990) argued that there are no practicable alternatives to it in evaluating research and proposed a checklist of requirements for any system of peer review, covering such matters as the choice of peers, transparency in its practice, and feedback on decisions. This also included the need to review practices continuously bearing in mind the inherent fallibility of peer review, to expose peer review practices thoroughly to the academic community, including providing information on peer review, and strategic and managerial input to decision-making, to choose peers well, to provide a return to the research community in the form of feedback from decisions, to demonstrate fully the equivalence in intellectual standards across all modes of support and to substantiate initial appraisal decisions by appropriate monitoring and ex-ante evaluation.

Meanwhile, in order to eradicate peer review dysfunction, Bozeman (1993) made some suggestions as follows;

- Guide No1: Use peer review in conjunction with other evaluation techniques
- Guide No2: Use peer review evaluation for R&D activities that are in the public domain
- Guide No3: Peers must be readily identifiable
- Guide No4.: Avoid internal peers
- Guide No5: Guard against dysfunctional group dynamics
- Guide No6: If scales are used, test the validity and reliability of those scales
- Guide No7: Provide a bias statement for reviewers

***Bibliometrics as the quantitative evaluation method***

Bibliometrics, or the study of publication-based data, is one of the more quantifiable methodological approaches available and has been largely embraced by researchers in the scientific community. As pressure for quantitative evidence of the success and impact of scientific efforts grows, the use of quantitative measures of science, such as bibliometrics, has become more pervasive. Its most common and accepted use is in the analysis of the output of basic research (Melkers, 1993). Bibliometrics is the study and analysis of scientific output using publication-based data. It is a clearly delineated body of research involving the measurement of physical units of publications, bibliographic citations, and surrogates for them (Broadus, 1987). Pritchard defined it as the application of mathematics and statistical methods to books and other media of communication (Pritchard, 1969). In practice, this means that the number of publications or citations generated by a research unit may be used to judge the productivity or output of that research group (Melkers, 1993). Publications serve to present in measurable terms the transfer of knowledge of R&D from one scientist to another.

Bibliometric approaches in assessing scientific output may be traced back to the turn of the century. However, formal bibliometric analysis is said to have originated in the 1960's. Derek de Solla Price and Eugene Garfield were the leaders in the movement to develop bibliometric indicators (Price, 1963; Garfield, et al., 1964). Henry Small helped to refine the method with the development of co-citation analysis (Small and Griffith, 1974). The creation of the Science Citation Index in 1961 allowed bibliometric analysis to become even more systematised (Garfield, 1979).

Although the term biblometric analysis is often used generally, it represents several different technical forms, each of which has different purposes and utilities. The common denominator in each of these forms is the concept that publications represent a flow of information primarily from basic science (Martin & Irvine, 1983). Each of these bibliometric approaches provide different information on the scientific enterprise. The main derivatives of bibliometrics are publication counts, citation counts, co-citation analysis, co-word analysis, scientific mapping (Melkers, 1993), patent analysis, citations in patents and co-nomination.

The development of bibliometrics greatly relies on how to make better use of the knowledge stored in the scientific information net such as journals and patents. As techniques in storing and analysing information are developed, bibliometrics is also becoming more complex and variously applied. Thus, it is being used to appraise the contributions of individual researchers, research groups or institutes and has become an important instrument in the reward and funding system of the science process world-wide. Furthermore, at the macro level, it has great significance as an instrument in the provision of advice on longer-term strategic analysis.

However, gaining acceptance for the use of bibliometrics as an appropriate quantitative and objective evaluation tool required much time and effort. There is many criticism of and hostility to it within the scientific community, based on a belief that it cannot adequately and satisfactorily reflect and consider the quality or competence of those evaluated, but only calculate their research productivity. In other words, the past research results can be easily measured as a quantitative indicator through some bibliometric techniques, such as citation analysis, but the current and future qualitative importance of past research achievements is more difficult to consider and evaluate precisely, using only bibliometric analysis. Actually, qualitative evaluation itself also has many problems. Therefore, in the evaluation process, constructive interaction and complementary interpretation between qualitative and quantitative measurements of research achievement and researchers themselves need to be introduced. As a consequence, many attempts have been made to achieve more accurate and objective evaluation in the measurement of research quality and quantity.

***Bibliometrics and its complementary relationship to the peer review system***

Until recently, peer review was the most common form of evaluation of research performance, but the application of bibliometric evaluation has now been proposed as an alternative or supplement to it.

Peer review systems in which decisions are heavily dependent on the selection of committee members, have some weaknesses in maintaining balanced and unbiased judgements in newly emerging or inter-disciplinary fields, where there are many conflicts of interest among priorities, due to the lack of science resources. Furthermore, they also have problems in correctly recognising and evaluating the quality of younger researchers or newcomers. Bibliometrics is useful as a supplementary tool for improving peer review-based evaluation, but also defective in that it cannot provide insight into most cognitive aspects of research work.

Both methods, although certainly not perfect, have their own considerable merits. Bibliometric methods can provide useful checkpoints and additional information to that obtained by conventional content-based peer review. Bibliometric analysis can never replace peer review judgements. On the other hand, peer review judgements alone cannot give sufficient information on important aspects of research productivity and on the impact of research activities. Therefore, peer review and bibliometric analysis will never be completely independent measures. However, an ideal combination of peer review and bibliometric analysis can certainly enrich the process of research evaluation in efficient and effective ways.

***Choice of experts as peer reviewers***

When peer review is used for evaluation, it usually relies on a few peers only. The choice of these experts is a decisive criterion of the validity of the evaluation. In order to conduct a successful evaluation, it is vital that peer members should be chosen on the basis of competence in their specialties which should cover whole targeted areas or fields, if possible. If industrial R&D is to be evaluated, some industrialists can also be included in the peer group. According to the circumstances, non-experts can be included in order to reflect broader opinions, for instance, when evaluating social impact. On the other hand, in some cases, it can be more effective to use foreign expert evaluators as external evaluative resources.

It is necessary that the selection of peer reviewers should take into account the overall research trends and orientation of all scientific fields or communities concerned as comprehensively as possible. A well balanced constitution of experts can greatly contribute to the impartiality of any judgement. A possible method to achieve such balance, during the

selection procedure, is to use a computerised system in order to eliminate such obstacles as too close a cooperative relationship between researchers and evaluators. In addition, establishing a standing evaluation organisation, such as a bank of specialties, which covers all the fields under consideration can save on time and costs. The involvement of the researchers in evaluation can gain agreement from them, because they may be the only experts in the relevant research field. Through contact between researchers and evaluators, useful and new scientific knowledge can be disseminated throughout the entire scientific community. However, this method requires a general consensus among scientists about the impartiality of evaluation.

On the other hand, in the case of problems arising from an insufficient pool of national research manpower in the given fields, for instance, in small countries such as Sweden etc, it is also possible to include international foreign scientists in the evaluation group. Using international experts can be positive and useful in giving the opportunity for the national researchers to see their discipline in a broader context. At the same time, however, it may cause some problems, because many mistakes and misunderstanding occur in communicating and interpreting a current national situation which may be impossible to scientifically evaluate. As an alternative approach, Georghiou (OECD, 1987) recommended seeking the views of expatriate scientists if they exist and are contactable in sufficient numbers.

## 4. Korean evaluation system

Since the early 1980s, Korea also has placed emphasis on strengthening competitive advantage through technological innovations (Kim, 1987; Kim and Dahlman, 1992) as is the case of other advanced countries. In line with this emphasis, the government initiated a series of national R&D programmes and formulated a long range development plan of science and technology towards the 2000s. (MOST, 1998) Simultaneously, the Korean government has paid increasing attention to evaluating R&D programmes (Lee et al, 1996).

With Korea's increasing concern about evaluating government R&D activities, the government has attempted formally to establish a national R&D evaluation function. Government and research institutions are recognising that more creativity and innovation in the research systems may be essential to future economic success, and are rapidly adopting and adapting Western evaluation concepts and techniques such as the use of external review committee and bibliometric data in a bid to improve the productivity and the quality of the research output.

However, overall Korean evaluation system has not yet been as firmly fixed as the Western

system, because there are still some issues for Korea to overcome in the foreseeable future. These include the lack of objectivity and credibility of the evaluation practices, insufficient expert evaluators, and the negative attitude toward evaluation arising from the oriental cultural difference with its reluctance to criticise other researchers' performance. Nevertheless, within a relatively short period, Korea successfully adopted and adapted the Western evaluation system. From now on, the main task will be to reestablish and reinforce Korea's own standardised system through careful review of some issues based on past experience, and this make it more suitable to Korean research environments.

In Korea, the most common evaluation method adopted and used in the relevant bodies including government is the peer review system. However, studies by many experts have revealed that this system has some problems of bias. The Korean peer review system is no exception to this. Since the subjectivity of peer review is inherent, it is best to supplement such a subjective evaluation with less subjective components, rather than seeking to reduce the subjectivity of peer review (Bozeman, 1993). Thus, it is necessary for Korea to use peer review in conjunction with less subjective evaluation methods.

As one alternative in order to avoid these problems of subjectivity of peer review, some bibliometric techniques have been introduced into Korean evaluation practices. It is considered that the impersonal nature of bibliometric methods has great merit in that it has undoubted appeal in a culture where people are loath to make direct personal evaluation of individuals. In practice, the publication counts technique has been the most frequently used of the various bibliometric methods in order to evaluate researchers' past performance. However, bibliometric approaches to evaluation give rise to other problems.

One good example of the shortcoming of over-emphasis on such data-based evaluation is that it encourages scientists to choose popular topics, instead of tackling valuable problems that make major contributions. Furthermore, people working in areas of hot international interest have higher chances of being cited. (Nature, 1997).

A more fundamental problem is that Korea has an insufficient pool of expert evaluators. To solve this problem, it is necessary to reconsider and reform the Korean university research and educational system because the university is represented and recognised as the leading actor in carrying out national R&D activities as well as in developing research manpower. The objectivity and credibility of the evaluation results will greatly depend on how expert evaluators are selected and chosen. Therefore, more emphasis should be placed on the university, which is an abundant pool of most of the high-qualified research manpower as well as the place with the greatest potential for research and the cultivation of young scientists in Korea.

Using foreign expert evaluators could be considered, but this is very difficult, because they do not usually have enough information on and understanding of Korean science and technology. In particular, such a method is absolutely impossible at the programme or project level, although such evaluators can be invited and utilised sometimes as advisors on the broad overall diagnosis for the entire system and future policy direction of R&D. Another method, although it too has similar problems, is the use of Korean expatriate scientists in evaluation practice, which is considered as the best alternative for solving the problem of the lack of sufficient experts in Korea. In some cases, they can give precious information about where current Korean S&T is positioned in the international arena and how the best future S&T development should be oriented.

Some problems may arise in administrative procedures in evaluation practices. Horizontal and vertical cooperation between ministries dealing with the national R&D programme, and between government and research supporting organisations is becoming a very important issue where there are conflicts of interests arising from organisational egoism.

Lastly, some of the inherent problems in Korean research environment have been further aggravating the impartiality and objectivity of evaluation results. One of these problems is an oriental cultural barrier against the Western evaluation system.

***Evaluation system***

Although the history of the introduction of the Western evaluation system into Korea is relatively short, the evaluation system has been modified and well operated until now. In Korea, the peer review system as a qualitative evaluation method has been predominantly used and, therefore, the decision by peer review has played a crucial role in evaluation practice in real terms. According to the definition by Bozeman (1993), the Korean peer review system can be classified as the traditional peer review where peer experts' judgements are the most important factor. There are two possible kinds of ways of trying to avoid impartiality and subjectivity, which are regarded as the most common problems in the peer review system. One possibility is to improve the functionality of the system itself and the other is to cultivate the expertise of the peers, which is the main factor in the improvement of the peer review system. Who the peers should be will be discussed in the following section. In this section, focus will be placed on how to improve the system itself.

In the practical terms, peer review based evaluation in Korea has been carried out mainly in the form of mail and sub committee and extended committee review. In some cases, self-appraisal and site-visit evaluation were recently introduced and incorporated into the national

R&D evaluation process. Self-appraisal is a useful auxiliary instrument because no one has as much knowledge about a project as the researchers themselves. However, they are liable to pursue self-interest. Site-visits can provide a rich source of information, but, as the reviewers are usually not prepared to devote more than 1 or 2 days to this, the contribution of site-visits is limited (Lee et al, 1996). The current peer review system should be modified and strengthened by the adoption of more objective criteria or factors in the evaluation process.

Therefore, approaches based on objective indicators, such as bibliometric approaches and socio-economic approaches, can be considered as a complementary tool of peer review in order to maintain the objectivity and credibility of evaluations. In particular, peer review is likely to be inappropriate for applied and commercialisation research where success is related to developing prototypes or commercialising technologies (Bozeman, 1993) because it is very difficult to identify appropriate peers. In such case, patent analyses can be supplemented in the evaluation process.

Bibliometric approaches which are initiated and developed in the Netherlands and UK, in order to combine objective factors with the peer review, can be regarded as an alternative modification of the evaluation system based on interpersonal decisions reflecting impersonalised evident information about the past, present and future possible research achievement and capability.

However, several studies revealed that bibliometrics also has many problems. For instance, the publication counting technique, which has been commonly used in the Korean evaluation process and as a science indicator to measure the ranking of a country's capability in research, does not distinguish the weighted factors between the quality and quantity of research. It just counts the number of published journals. It is also difficult to evaluate precisely the actual achievement of Korean researchers, because most scientific journals are published in English only.

Therefore, a prerequisite for the successful use of bibliometric techniques is the invention of a mechanism which can consider both qualitative and quantitative factors of the scientific articles published in the journals. The comparative importance of articles written in Korean and English should also be considered. To do this, overall investigation of the scientific journals in Korea should be carried out, together with deeper study of the methodologies developed and used by ISI in sorting the journals by citation impacts etc.

Also, evaluation takes place mainly in selection procedures of programme and projects in Korea. Although interim and *ex-post* evaluation have also been carried out, the basic purpose

of them has not been satisfactory. Evidence can be addressed in that there are rarely programmes or projects which are modified and terminated as a result of the evaluation result. Once selected, programme and projects used to be funded to the end, regardless of interim evaluation. Also, the result of ex-post evaluation did not apply to the next selection procedures and policy-making process. Therefore, one of the suggestions is that interim evaluation should be carried out in an effective way in which the progress of the programme and projects should be precisely checked in the middle stage and the result of interim evaluation should be reflected in the next stage funding decision as well as in the initial stage planning again. Also the result of *ex-post* evaluation should be fed back into the planning and policy-making process.

***Selection and cultivation of appropriate expert peers***

The identification and selection of appropriate reviewers should be an important success factor for the peer review method. However, it is difficult to invite appropriate reviewers because of the limited number of qualified experts in Korea. Furthermore, in this situation where expert evaluators are insufficient, the only solution for improving and cultivating the expertise of evaluators should be sought within the university system, because the university is a pool of high-qualified manpower in Korea. However, universities themselves also have some problems as provider of evaluators and expertise. One of them is that there is too strong an invisible network between researchers through their having the same academic career, in particular, in high school and university and through coming from the same region. In evaluation practice, efforts should be constantly made to exclude this unnecessary relationship between researchers in selecting evaluators. To do so, the collection of information regarding who works in what area etc is needed by establishing a data base. Also, reviewers should be constantly asked to be impartial.

In the longer term, it will also be important to cultivate promising young scientists as future expert evaluators. It is necessary to support them through various programme, so that they may have the capability to carry out their own research by themselves.

Peer review in the form of committee meetings is likely to experience dysfunctional group dynamics, such as group think, a bandwagon effect and domination by more vocal participants. To guard against these dynamics, it is useful to have a peer reviewer provide a separate, individual assessment before beginning discussions (Bozeman, 1993). Also, it is preferable to adjust the composition of reviewers for the different project types (Lee et al, 1994). It is possible to evaluate the reviewers' reliability and performance, if the evaluation activities of reviewers are well recorded. Through this, unreliable reviewers can be identified and excluded from future evaluations. Also, it will be helpful to provide reviewers with an

explanation about personal bias and potential conflicts occurring in the peer review process. While such a statement hardly addresses all the possible pitfalls of biased peer review, it can provide a partial remedy (Bozeman, 1993).

To compensate for the lack of expert evaluators, it is possible to invite foreign experts as in Sweden, even though the language barrier should be lifted in advance. Such language problem is hard to overcome in short term. Therefore, the effective use of expatriate Korean scientists can be considered when using foreign experts in evaluation group or committees. However, their use should be confined to the evaluation of overall national R&D activities and policy, rather than of project selection. It is also important to give sufficient information to them about the researchers and proposals evaluated. In longer term, it is necessary to cultivate expert evaluators by the active exchange of scientists with other countries and overseas training of young students.

***Reform of administrative procedures***

In the evaluation process, the role of administrators as intermediate agents between the evaluators and researchers is very important. In particular, in pursuit of the national R&D activities, close cooperation and full exchange of the information between the relevant bodies within government is recognised as a crucial factor for success. In other words, various R&D programmes initiated and funded by several ministries and supporting sub-organisations should be well coordinated through cooperation between them. R&D has become more interdisciplinary and related to more branches of government. Cooperation among the ministries seems to be particularly critical and indispensable for the efficient use of the national resources. This linkage is also needed for the ministries to share the methods and experiences in the R&D programme management and evaluation.

Tight vertical linkage between a ministry and the responsible organisation is important, but this linkage should not impinge on the autonomy of the latter. As a reminder of the rationale, the division of roles is necessary when the government should formulate national R&D policies, and the responsible organisations should conduct the professional tasks of managing and evaluating the national R&D programmes (Lee et al, 1996).

In addition, the setting up of a formal format regarding evaluation procedures is necessary. As in the case of the UK ROAME, a clear statement about the overall purpose and procedures etc. of evaluation should be established, so that all evaluation activities can follow this guideline. Evaluation criteria and their weightings should be appropriately chosen and adjusted for the changes of R&D environment and for the types of R&D activities. They should also be made accessible to researchers prior to the start of programme and projects. Evaluation results

should also be given to the researchers.

In order to adopt and adapt the objective factors in the evaluation process based on peer review in an effective way, a great deal of effort is required in collecting, analysing and intergrating the various data and giving evaluation results so that they are timely for the policy making process.

## 8. Conclusion

The effects or results of evaluation can be regarded as positive from the point of view that they may bring about constructive changes, reforms and improvement of research quality. However, if too much attention is paid to quality and efficiency, this may lead to too competitive an atmosphere which may impede scientists from actively co-operating in research. Furthermore, too much emphasis only on products and efficiency may result in short-term policies that do not take enough risk. The balance between evaluation and the development of science and technology should be considered in terms of the limited resources and their efficient use.

Another very crucial factor is how to maintain the transparency or impartiality of evaluation practices. In this process, precise interpretation of the evaluation results is important. Sometimes, bibliometric analysis can be used as a substitute for qualitative evaluation in order to gain objective evaluation results. At the final stage, all evaluation results which have detected problems in the relevant research activities at any level should be reflected directly in the policy-making process. Through this circular process involving the perception of problems, their analysis and the reflection of them in the policy-making process, and then newly setting up final aims and their implementation, evaluation and policy-making have an indivisible reciprocal relationship.

Evaluations are not a passing fashion in science policy, but apparently a policy tool that has come to stay. They are necessitated by the growing costs and scope of scientific research activities and increasing demands for accountability for the use of public funds. There does not exist a standardised set of evaluation methods which can be used in a routine manner, but all evaluations have to be prepared with care and skill, with methods selected for the respective purposes. An important prerequisite for the use of evaluations is an adequate follow-up mechanism (Luukkonen-Gronow, 1987).

Evaluation depends on the specific function of science and technology, which differs from country to country, according to the peculiar structure and organisation of the national R&D system. This implies that evaluation will not be perceived and approached in the same way in

every country. Furthermore, the effects of the output of evaluation are greatly dependent on where the emphasis will be placed for the advancement of science and technology and for the efficient allocation and use of limited resources.

As for evaluation methods, the peer review system is considered the most pervasive one so far in the scientific community even though it has some problems as already outlined. In a bid to remove its limitations, some derivations, such as bibliometrics, have been developed as a complementary mechanism. However, while only the development of more refined techniques of evaluation can help increase its efficiency to some extent, at the end of the day, this will not result in a perfect process of evaluation. Therefore, further study should concentrate more on the efficient management of the evaluation exercise and adequate use or application of the results of evaluation.

In addition, one of main factors for the successful evaluation will be how to select and, in longer-term, also to cultivate expert-evaluators. More efforts should be concentrated on to upgrade the university research and its manpower capabilities in close cooperation with other major R&D bodies such as industry and governmental research institutes.

On the other hand, reforming administrative procedures is very important in order to make the whole evaluation process more creative and productive. Close cooperation between the relevant government supporting bodies also needs to be strengthened. Securing transparency in operating all evaluative procedures can also help improve the impartiality of evaluation itself.

In particular, Western style evaluation methods and techniques developed until now to meet the needs of some leading developed countries, when applied to oriental countries, are unlikely to be clearly understood and practised because of particular oriental traditions and cultures, such as the strong reluctance to criticise colleagues or researchers, a strong fundamental resistance to evaluation, in particular, at the individual level and the general recognition of the Confucian and Buddhist values of respect and group harmony. In this regard, there is a need, perhaps, to modify these traditional evaluation models to render them more suitable to the cultural Korea's scientific environment.

*References*

**Advisory Board for the Research Councils (1990)** "Peer review (the Boden report)," London: ABRC

**Bozeman, B. (1993)** " Peer Review and Evaluation of R&D Impacts," in Barry Bozeman and Julia Melkers (eds.) Evaluating R&D Impacts: Methods and Practice, Boston: Kluwer Academic Publishers.

**Broadus, R.N. (1987)** "Toward a definition of bibliometrics," Scientometrics (12) p373-379

Cabinet Office (1989) "R&D assessment: A guide for customers and managers of research and development," London: HMSO

**Chubin, D.E. and Hackett, E.J. (1990)** "Peerless science: Peer review and U.S. science policy," New York: State University of New York Press.

**Chubin, D.E. and Jasanoff, S. (1985)** "Peer review and public policy," Science, Technology, Human values (10) p3-5

**Freeman, H. E. (1977)** "The present status of evaluation research," in Guttentag, M.A. and Saar, S. (eds.) Evaluation studies review annual (2) p17-51

**Garfield, E. (1979)** "Is citation analysis a legitimate evaluation tool?," Scientometrics (1) p359-375

**Georghiou, L. (1989)** " Organization of Evaluation," in Ciba Foundation (ed.) The Evaluation of Scientific Research, Chichester: John Wiley & Sons.

**Kim, L. and Dahlman, C.J. (1992)** "Technology policy for industialisation: An integrative framework and Korea's experience," Research policy (21) p437-452

**Kim, Y.W. (1987)** "R&D activities in private sector and promotion policy," in proceedings of the workshop on the formulation of science and technology policy in developing countries, Seoul: MOST

**King J. (1987)** "A review of bibliometric and other science indicators and their role in research evaluation," Journal of information science 13 p261-276

**Kostoff, R.N. (1994)** "Assessing research impact: Federal peer review practices," Evaluation review 18(1) p31-40

**Kruytbosch, Carlos (1989)** "The role and effectiveness of peer review" in Ciba Foundation (ed.) The Evaluation of Scientific Research, Chichester: John Wiley & Sons.

**Lee, M. Son, B. and Jan, H. (1994)** " A study on the an effective evaluation system for national R&D projects," Korean Management Review (23) p193-227 (in Korean)

**Lee, M., Son, D. and Om, K. (1996)** "Evaluation of national R&D projects in Korea," Research policy (25) p805-818

**Luukkonen-Gronow, T. (1987)** " Scientific Research Evaluation: A review of methods and various contexts of their application," R&D Management 17(3) p207-221.

**Martin, B. R. and Irvine, J. (1983)** "Assessing basic research: Some partial indicators of scientific progress in radio astronomy", Research Policy (12) p61-90

**Melkers Julia (1993)** "Bibliometrics as a tool for analysis of R&D impact" in Barry Bozeman and Julia Melkers (eds.) Evaluating R&D Impacts: Methods and Practice, Boston : Kluwer Academic Publishers.

**MOST (1998)** "Science and Technology in Korea," Seoul: MOST

**Nature (1997)** " Western research assessment meets Asian cultures," Nature (389) p113-117

**OECD (1987)** "Evaluation of research: A selection of current practices", Paris: OECD.

**OECD (1990)** "Evaluation of programmes: Promoting research, technology and innovation", Paris: OECD.

**Patton, M. Q. (1986)** "Utilisation-focused evaluation," London: Sage publications

**Price, D. de solla (1963)** "Big science, little science," New York: Columbia University Press

**Pritchard, A(1969)** "Statistical bibliography or bibliometrics," Journal of documentation (25) p358-359

**Rossi, P. H. and Freeman, H.E.(1989)** "Evaluation: A systematic approach, 4th Edition," London: Sage publications

**Small, H. and Griffith, B. (1974)** "The structure of scientific literatures," Science studuies (4) p17-40

**Weinberg, A (1984)** "Values in science: Unity as a criterion of scientific choice", Minerva 22(1) p1-12

**Westerheijden, D. (1991)** "Promises, problems and pitfalls of peer review: The use of peer review in external quality assessment in higher education," presented at the third international seminar on assessing quality in higher education, Bath, 16-19 July 1991

**Zuckerman, H. and Merton, R. (1973)** "Institutionalised patterns of evaluation in science" in Merton, R. (ed.) The sociology of science, Chicago: University of chicago press.